

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12P 7/62, C12N 9/88, 15/63, C07H 21/04		A1	(11) International Publication Number: WO 97/22711 (43) International Publication Date: 26 June 1997 (26.06.97)
(21) International Application Number: PCT/US96/20119		(81) Designated States: CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
(22) International Filing Date: 18 December 1996 (18.12.96)		Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(30) Priority Data: 60/008,847 19 December 1995 (19.12.95) US			
(71) Applicant (for all designated States except US): REGENTS OF THE UNIVERSITY OF MINNESOTA [US/US]; Morrill Hall, 100 Church Street S.E., Minneapolis, MN 55455 (US).			
(72) Inventors; and			
(75) Inventors/Applicants (for US only): SHERMAN, David, H. [US/US]; 2248 Drew Avenue South, St. Louis Park, MN 55416 (US). WILLIAMS, Mark, D. [US/US]; 1414 Breda Avenue, St. Paul, MN 55108 (US). XUE, Yongquan [CN/US]; 1403 Cleveland Avenue North, St. Paul, MN 55108 (US).			
(74) Agent: VIKSNINS, Ann, S.; Schwegman, Lundberg, Woessner & Kluth, P.O. Box 2938, Minneapolis, MN 55402 (US).			

(54) Title: METABOLIC ENGINEERING OF POLYHYDROXYALKANOATE MONOMER SYNTHASES

(57) Abstract

A novel pathway for the synthesis of polyhydroxyalkanoates is provided. A method of synthesizing a recombinant polyhydroxyalkanoate monomer synthase is also provided. These recombinant polyhydroxyalkanoate synthases are derived from multifunctional fatty acid synthases or polyketide synthases and generate hydroxyacyl acids capable of polymerization by a polyhydroxyalkanoate synthase.

BEST AVAILABLE COPY

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LJ	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	I.U	Luxembourg	TD	Chad
CZ	Czech Republic	I.V	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

METABOLIC ENGINEERING OF POLY-HYDROXYALKANOATE MONOMER SYNTHASES

5

Background of the Invention

Polyhydroxyalkanoates (PHAs) are one class of biodegradable polymers. The first identified member of the PHAs thermoplastics was polyhydroxybutyrate (PHB), the polymeric ester of 10 D(-)-3-hydroxybutyrate.

The biosynthetic pathway of PHB in the gram negative bacterium *Alcaligenes eutrophus* is depicted in Figure 1. PHAs related to PHB differ in the structure of the pendant arm, R (Figure 2). For example, R=CH₃ in PHB, while R=CH₂C H₃ in 15 polyhydroxyvalerate, and R=(CH₂)₄CH₃ in polyhydroxyoctanoate.

The genes responsible for PHB synthesis in *A. eutrophus* have been cloned and sequenced. (Peoples et al., *J. Biol. Chem.*, **264**, 15293 (1989); Peoples et al., *J. Biol. Chem.*, **264**, 15298 (1989)). Three enzymes: β -ketothiolase (*phbA*), acetoacetyl-CoA reductase (*phbB*), 20 and PHB synthase (*phbC*) are involved in the conversion of acetyl-CoA to PHB. The PHB synthase gene encodes a protein of M_r=63,900 which is active when introduced into *E. coli* (Peoples et al., *J. Biol. Chem.*, **264**, 15298 (1989)).

Although PHB represents the archetypical form of a 25 biodegradable thermoplastic, its physical properties preclude significant use of the homopolymer form. Pure PHB is highly crystalline and, thus, very brittle. However, unique physical properties resulting from the structural characteristics of the R groups in a PHA copolymer may result in a polymer with more desirable 30 characteristics. These characteristics include altered crystallinity, UV weathering resistance, glass to rubber transition temperature (T_g), melting temperature of the crystalline phase, rigidity and durability (Holmes et al., EPO 00052 459; Anderson et al., *Microbiol. Rev.*, **54**, 450 (1990)). Thus, these polyesters behave as thermoplastics, with melting

temperatures of 50-180°C, which can be processed by conventional extension and molding equipment.

Traditional strategies for producing random PHA copolymers involve feeding short and long chain fatty acid monomers to bacterial cultures. However, this technology is limited by the monomer units which can be incorporated into a polymer by the endogenous PHA synthase and the expense of manufacturing PHAs by existing fermentation methods (Haywood et al., FEMS Microbiol. Lett., 57, 1 (1989); Poi et al., Int. J. Biol. Macromol., 12, 106 (1990); Steinbuchel et al., In: Novel Biomaterials from Biological Sources. D. Byron (ed.), MacMillan, NY (1991); Valentin et al., Appl. Microbiol. Biotechnical, 36, 507 (1992)).

The production of diverse hydroxyacylCoA monomers for homo- and co-polymeric PHAs also occurs in some bacteria through the reduction and condensation pathway of fatty acids. This pathway employs a fatty acid synthase (FAS) which condenses malonate and acetate. The resulting β-keto group undergoes three processing steps, β-keto reduction, dehydration, and enoyl reduction, to yield a fully saturated butyryl unit. However, this pathway provides only a limited array of PHA monomers which vary in alkyl chain length but not in the degree of alkyl group branching, saturation, or functionalization along the acyl chain.

The biosynthesis of polyketides, such as erythromycin, is mechanistically related to formation of long-chain fatty acids. However, polyketides, in contrast to FASs, retain ketone, hydroxyl, or olefinic functions and contain methyl or ethyl side groups interspersed along an acyl chain comparable in length to that of common fatty acids. This asymmetry in structure implies that the polyketide synthase (PKS), the enzyme system responsible for formation of these molecules, although mechanistically related to a

FAS, results in an end product that is structurally very different than that of a long chain fatty acid.

Because PHAs are biodegradable polymers that have the versatility to replace petrochemical-based thermoplastics, it is 5 desirable that new, more economic methods be provided for the production of defined PHAs. Thus, what is needed are methods to produce recombinant PHA monomer synthases for the generation of PHA polymers.

10

Summary of the Invention

The present invention provides a method of preparing a polyhydroxyalkanoate synthase. The method comprises introducing an expression cassette into a non-plant eukaryotic cell. The expression cassette comprises a DNA molecule encoding a 15 polyhydroxyalkanoate synthase operably linked to a promoter functional in the non-plant eukaryotic cell. The DNA molecule encoding the polyhydroxyalkanoate synthase is then expressed in the cell. Thus, another embodiment of the invention provides a purified, isolated recombinant polyhydroxybutyrate synthase.

20

Another embodiment of the invention is a method of preparing a polyhydroxyalkanoate polymer. The method comprises introducing a first expression cassette and a second expression cassette into a eukaryotic cell. The first expression cassette comprises a DNA segment encoding a fatty acid synthase in which the dehydrase 25 activity has been inactivated that is operably linked to a promoter functional in the eukaryotic cell. The second expression cassette comprises a DNA segment encoding a polyhydroxyalkanoate synthase operably linked to a promoter functional in the eukaryotic cell. The DNA segments in the expression cassettes are expressed in the cell so 30 as to yield a polyhydroxyalkanoate polymer.

Another embodiment of the invention is a baculovirus expression cassette comprising a nucleic acid molecule encoding a polyhydroxyalkanoate synthase operably linked to a promoter functional in an insect cell.

5 The present invention also provides an expression cassette comprising a nucleic acid molecule encoding a polyhydroxyalkanoate monomer synthase operably linked to a promoter functional in a host cell. The nucleic acid molecule comprises a plurality of DNA segments. Thus, the nucleic acid
10 molecule comprises at least a first and a second DNA segment. No more than one DNA segment is derived from the *eryA* gene cluster of *Saccharopolyspora erythraea*. The first DNA segment encodes a first module and the second DNA segment encodes a second module, wherein the DNA segments together encode a polyhydroxyalkanoate
15 synthase.

Also provided is an isolated and purified DNA molecule. The DNA molecule comprises a plurality of DNA segments. Thus, the DNA molecule comprises at least a first and a second DNA segment. The first DNA segment encodes a first module and the
20 second DNA segment encodes a second module. No more than one DNA segment is derived from the *eryA* gene cluster of *Saccharopolyspora erythraea*. Together the DNA segments encode a recombinant polyhydroxyalkanoate monomer synthase. A preferred embodiment of the invention employs a first DNA segment derived
25 from the *vep* gene cluster of *Streptomyces*. Another preferred embodiment of the invention employs a second DNA segment derived from the *tyl* gene cluster of *Streptomyces*.

Yet another embodiment of the invention is a method of providing a polyhydroxyalkanoate monomer. The method comprises
30 introducing a DNA molecule into a host cell. The DNA molecule comprises a DNA segment encoding a recombinant

polyhydroxyalkanoate monomer synthase operably linked to a promoter functional in the host cell. The DNA encoding the recombinant polyhydroxyalkanoate monomer synthase, which synthase comprises at least a first module and a second module, is
5 expressed in the host cell so as to generate a polyhydroxyalkanoate monomer.

Also provided is a method of preparing a polyhydroxyalkanoate polymer. The method comprises introducing a first DNA molecule and a second DNA molecule into a host cell. The
10 first DNA molecule comprises a DNA segment encoding a recombinant polyhydroxyalkanoate monomer synthase. The recombinant polyhydroxyalkanoate monomer synthase comprises a plurality of modules. Thus, the monomer synthase comprises at least a first module and a second module. The first DNA molecule is
15 operably linked to a promoter functional in a host cell. The second DNA molecule comprises a DNA segment encoding a polyhydroxyalkanoate synthase operably linked to a promoter functional in the host cell. The DNAs encoding the recombinant polyhydroxyalkanoate monomer synthase and polyhydroxyalkanoate
20 synthase are expressed in the host cell so as to generate a polyhydroxyalkanoate polymer.

Yet another embodiment of the invention is an isolated and purified DNA molecule. The DNA molecule comprises a plurality of DNA segments. That is, the DNA molecule comprises at
25 least a first and a second DNA segment. The first DNA segment encodes a fatty acid synthase and the second DNA segment encodes a module of a polyketide synthase. A preferred embodiment of the invention employs a second DNA segment encoding a module which comprises a β -ketoacyl synthase amino-terminal to an acyltransferase

which is amino-terminal to a ketoreductase which is amino-terminal to an acyl carrier protein which is amino-terminal to a thioesterase.

The invention also provides a method of preparing a polyhydroxyalkanoate monomer. The method comprises introducing 5 a DNA molecule comprising a plurality of DNA segments into a host cell. Thus, the DNA molecule comprises at least a first and a second DNA segment. The first DNA segment encodes a fatty acid synthase operably linked to a promoter functional in the host cell. The second DNA segment encodes a polyketide synthase. The second DNA 10 segment is located 3' to the first DNA segment. The first DNA segment is linked to the second DNA segment so that the encoded protein is expressed as a fusion protein. The DNA molecule is then expressed in the host cell so as to generate a polyhydroxyalkanoate monomer.

15 Another embodiment of the invention is an expression cassette comprising a DNA molecule comprising a DNA segment encoding a fatty acid synthase and a polyhydroxyalkanoate synthase.

Also provided is a method of providing a polyhydroxyalkanoate monomer synthase. The method comprises 20 introducing an expression cassette into a host cell. The expression cassette comprises a DNA molecule encoding a polyhydroxyalkanoate monomer synthase operably linked to a promoter functional in the host cell. The monomer synthase comprises a plurality of modules. Thus, the monomer synthase comprises at least a first and second 25 module which together encode the monomer synthase.

A further embodiment of the invention is an isolated and purified DNA molecule comprising a DNA segment which encodes a *Streptomyces venezuelae* polyhydroxyalkanoate monomer synthase, a biologically active variant or subunit thereof. Preferably, the DNA 30 segment encodes a polypeptide having an amino acid sequence comprising SEQ ID NO:2. Preferably, the DNA segment comprises

SEQ ID NO:1. The DNA molecules of the invention are double stranded or single stranded. A preferred embodiment of the invention is a DNA molecule that has at least about 70%, more preferably at least about 80%, and even more preferably at least about 5 90%, identity to the DNA segment comprising SEQ ID NO:1, e.g., a "variant" DNA molecule. A variant DNA molecule of the invention can be prepared by methods well known to the art, including oligonucleotide-mediated mutagenesis. See Adelman et al., DNA, 2, 183 (1983) and Sambrook et al., Molecular Cloning: A Laboratory 10 Manual (1989).

The invention also provides an isolated, purified polyhydroxyalkanoate monomer synthase, e.g., a polypeptide having an amino acid sequence comprising SEQ ID NO:2, a biologically active subunit, or a biologically active variant thereof. Thus, the invention 15 provides a variant polypeptide having at least about 80%, more preferably at least about 90%, and even more preferably at least about 95%, identity to the polypeptide having an amino acid sequence comprising SEQ ID NO:2. A preferred variant polypeptide, or subunit of a polypeptide, of the invention includes a variant or subunit 20 polypeptide having at least about 10%, more preferably at least about 50% and even more preferably at least about 90%, the activity of the polypeptide having the amino acid sequence comprising SEQ ID NO:2. Preferably, a variant polypeptide of the invention has one or 25 more conservative amino acid substitutions relative to the polypeptide having the amino acid sequence comprising SEQ ID NO:2. For example, conservative substitutions include aspartic-glutamic as acidic amino acids; lysine/arginine/histidine as basic amino acids; leucine/isoleucine, methionine/valine, alanine/valine as hydrophobic amino acids; serine/glycine/alanine/threonine as 30 hydrophilic amino acids. The biological activity of a polypeptide of the invention can be measured by methods well known to the art.

As used herein, a "linker region" is an amino acid sequence present in a multifunctional protein which is less well conserved in amino acid sequence than an amino acid sequence with catalytic activity.

5 As used herein, an "extender unit" catalytic or enzymatic domain is an acyl transferase in a module that catalyzes chain elongation by adding 2-4 carbon units to an acyl chain and is located carboxy-terminal to another acyl transferase. For example, an extender unit with methylmalonylCoA specificity adds acyl groups to
10 a methylmalonylCoA molecule.

As used herein, a "polyhydroxyalkanoate" or "PHA" polymer includes, but is not limited to, linked units of related, preferably heterologous, hydroxyalkanoates such as 3-hydroxybutyrate, 3-hydroxyvalerate, 3-hydroxycaproate, 3-hydroxyheptanoate, 3-hydroxyhexanoate, 3-hydroxyoctanoate, 3-hydroxyundecanoate, and 3-hydroxydodecanoate, and their 4-hydroxy and 5-hydroxy counterparts.
15

As used herein, a "Type I polyketide synthase" is a single polypeptide with a single set of iteratively used active sites. This is in contrast to a Type II polyketide synthase which employs active sites on
20 a series of polypeptides.

As used herein, a "recombinant" nucleic acid or protein molecule is a molecule where the nucleic acid molecule which encodes the protein has been modified *in vitro*, so that its sequence is
25 not naturally occurring, or corresponds to naturally occurring sequences that are not positioned as they would be positioned in a genome which has not been modified.

As used herein, a "multifunctional protein" is one where two or more enzymatic activities are present on a single polypeptide.

As used herein, a "module" is one of a series of repeated units in a multifunctional protein, such as a Type I polyketide synthase or a fatty acid synthase.

As used herein, a "premature termination product" is a 5 product which is produced by a recombinant multifunctional protein which is different than the product produced by the non-recombinant multifunctional protein. In general, the product produced by the recombinant multifunctional protein has fewer acyl groups.

As used herein, a DNA that is "derived from" a gene 10 cluster, is a DNA that has been isolated and purified *in vitro* from genomic DNA, or synthetically prepared on the basis of the sequence of genomic DNA.

As used herein, the terms "isolated and/or purified" refer to *in vitro* isolation of a DNA or polypeptide molecule from its 15 natural cellular environment, and from association with other components of the cell, such as nucleic acid or polypeptide, so that it can be sequenced, replicated and/or expressed. Moreover, the DNA may encode more than one recombinant Type I polyketide synthase and/or fatty acid synthase. For example, "an isolated DNA molecule 20 encoding a polyhydroxyalkanoate monomer synthase" is RNA or DNA containing greater than 7, preferably 15, and more preferably 20 or more sequential nucleotide bases that encode a biologically active polypeptide, fragment, or variant thereof, that is complementary to the non-coding, or complementary to the coding strand, of a 25 polyhydroxyalkanoate monomer synthase RNA, or hybridizes to the RNA or DNA encoding the polyhydroxyalkanoate monomer synthase and remains stably bound under stringent conditions, as defined by methods well known to the art, e.g., in Sambrook et al., *supra*.

Brief Description of the Figures

Figure 1: The PHB biosynthetic pathway in *A. eutrophus*.

Figure 2: Molecular structure of common bacterial
5 PHAs. Most of the known PHAs are polymers of 3-hydroxy acids
possessing the general formula shown. For example, R=CH₃ in PHB,
R=CH₂CH₃ in polyhydroxyvalerate (PHV), and R=(CH₂)₄CH₃ in
polyhydroxyoctanoate (PHO).

Figure 3: Comparison of the natural and
10 recombinant pathways for PHB synthesis. The three enzymatic steps
of PHB synthesis in bacteria involving 3-ketothiolase, acetoacetyl-CoA
reductase, and PHB synthase are shown on the left. The two
enzymatic steps involved in PHB synthesis in the pathway in *Sf21*
cells containing a rat fatty acid synthase with an inactivated dehydrase
15 domain (ratFAS206) are shown on the right.

Figure 4: Schematic diagram of the molecular
organization of the *tyl* polyketide synthase (PKS) gene cluster. Open
arrows correspond to individual open reading frames (ORFs) and
numbers above an ORF denote a multifunctional module or synthase
20 unit (SU). AT=acyltransferase; ACP=acyl carrier protein; KS=β-
ketoacyl synthase; KR=ketoreductase; DH=dehydrase; ER=enoyl
reductase; TE=thioesterase; MM=methylmalonylCoA; M=malonyl
CoA; EM=ethylmalonyl CoA. Module 7 in *tyl* is also known as
Module F.

25 Figure 5: Schematic diagram of the molecular
organization of the *met* PKS gene cluster.

Figure 6: Strategy for producing a recombinant PHA
monomer synthase by domain replacement.

Figure 7: (A) 10% SDS-PAGE gel showing samples
30 from various stages of the purification of PHA synthase; lane 1,

molecular weight markers; lane 2, total protein of uninfected insect cells; lane 3, total protein of insect cells expressing a rat FAS (200 kDa; Joshi et al., *Biochem J.* **296**, 143 (1993)); lane 4, total protein of insect cells expressing PHA synthase; lane 5, soluble protein from sample in 5 lane 4; lane 6, pooled hydroxylapatite (HA) fractions containing PHA synthase. (B) Western analysis of an identical gel using rabbit- α -PHA synthase antibody as probe. Bands designated with arrows are: a, intact PHB synthase with N-terminal alanine at residue 7 and serine at residue 10 (A7/S10); b, 44 kDa fragment of PHB synthase with N- 10 terminal alanine at residue 181 and asparagine at residue 185 (A181/N185); c, PHB synthase fragment of approximately 30 kDa apparently blocked based on resistance to Edman degradation; d, 22 kDa fragment with N-terminal glycine at residue 187 (G187). Band d apparently does not react with rabbit- α -PHB synthase antibody (B, lane 15 6). The band of similar size in B, lane 4 was not further identified.

Figure 8: N-terminal analysis of PHA synthase purified from insect cells. (a) The expected N-terminal 25 amino acid sequence of *A. eutrophus* PHA synthase. (b&c) The two N-terminal sequences determined for the *A. eutrophus* PHA synthase produced 20 in insect cells. The bolded sequences are the actual N-termini determined.

Figure 9: Spectrophotometric scans of substrate, 3-hydroxybutyrate CoA (HBCoA) and product, CoA. The wavelength at which the direct spectrophotometric assays were carried out (232 nm) 25 is denoted by the arrow; substrate, HBCoA (•) and product, CoA (◦).

Figure 10: Velocity of the hydrolysis of HBCoA as a function of substrate concentration. Assays were carried out in 40 or 200 μ l assay volumes with enzyme concentration remaining constant at 0.95 mg/ml (3.8 μ g/40 μ l assay). Velocities were calculated from the 30 linear portions of the assay curves subsequent to the characteristic lag

period. The substrate concentration at half-optimal velocity, the apparent K_m value, was estimated to be 2.5 mM from this data.

Figure 11: Double reciprocal plot of velocity versus substrate concentration. The concave upward shape of this plot is 5 similar to results obtained by Fukui et al. (Arch. Microbiol., **110**, 149 (1976)) with granular PHA synthase from *Z. ramigera*.

Figure 12: Velocity of the hydrolysis of HBCoA as a function of enzyme concentration. Assays were carried out in 40 μ l assay volumes with the concentration HBCoA remaining constant at 10 8 μ M.

Figure 13: Specific activity of PHA synthase as a function of enzyme concentration.

Figure 14: pH activity curve for soluble PHA synthase produced using the baculovirus system. Reactions were carried out in 15 the presence of 200 mM P_i . Buffers of pH<10 were prepared with potassium phosphate, while buffers of pH>10 were prepared with the appropriate proportion of Na_3PO_4 .

Figure 15: Assays of the hydrolysis of HBCoA with varying amounts of PHA synthase. Assays were carried out in 40 μ l 20 assay volumes with the concentration of HBCoA remaining constant at 8 μ M. Initial A_{232} values, originally between 0.62 and 0.77, were normalized to 0.70. Enzyme amounts used in these assays were, from the upper-most curve, 0.38, 0.76, 1.14, 1.52, 1.90, 2.28, 2.66, 3.02, 3.42, 7.6, and 15.2 μ g, respectively.

Figure 16: SDS/PAGE analysis of proteins synthesized 25 at various time-points during infection of *Sf21* cells. Approximately 0.5 mg of total cellular protein from various samples was fractionated on a 10% polyacrylamide gel. Samples include: uninfected cells, lanes 1-4, days 0, 1, 2, 3 respectively; infection with BacPAK6::phbC alone,

lanes 5-8, days 0, 1, 2, 3 respectively; infection with baculoviral clone containing ratFAS206 alone, lanes 9-12, days 0, 1, 2, 3 respectively; and ratFAS206 and BacPAK6 infected cells, lanes 13-16 days 0, 1, 2, 3, respectively. A=mobility of FAS, B=mobility of PHA synthase.

5 Molecular weight standard lanes are marked M.

Figure 17: Gas chromatographic evidence for PHB accumulation in *Sf21* cells. Gas chromatograms from various samples are superimposed. PHB standard (Sigma) is chromatogram #7 showing a propylhydroxybutyrate elution time of 10.043 minutes
10 (s, arrow). The gas chromatograms of extracts of the uninfected (#1); singly infected with ratFAS206 (#2, day 3); and singly infected with PHA synthase (#3, day 3) are shown at the bottom of the figure. Gas chromatograms of extracts of dual-infected cells at day 1 (#4), 2 (#5), and 3 (#6) are also shown exhibiting a peak eluting at 10.096 minutes
15 (x, arrow). The peak of dual-infected, day 3 extract (#6) was used for mass spectrometry (MS) analysis.

Figure 18: Gas-chromatography-mass spectrometry analysis of PHB. The characteristic fragmentation of propylhydroxybutyrate at m/z of 43, 60, 87, and 131 is shown. A) standard PHB from bacteria (Sigma), and B) peak X from ratFAS206 and BacPAK6: phbC baculovirus infected, day 3 (#6, Figure 17) *Sf21* cells expressing rat FAS dehydrase inactivated protein and PHA synthase.

Figure 19: Map of the *vep* (*Streptomyces venezuelae* polyene encoding) gene cluster.
25

Figure 20: Plasmid map of pDHS502.

Figure 21: Plasmid map of pDHS505.

Figure 22: Cloning protocol for pDHS505.

Figure 23: Nucleotide sequence (SEQ ID NO:1) and corresponding amino acid sequence (SEQ ID NO:22) of the *vep* ORFI.
30

Detailed Description of the Invention

The invention described herein can be used for the production of a diverse range of biodegradable PHA polymers through genetic redesign of DNA encoding a FAS or *Streptomyces* spp. Type I PKS polypeptide to provide a recombinant PHA monomer synthase. Different PHA synthases can then be tested for their ability to polymerize the monomers produced by the recombinant PHA synthase into a biodegradable polymer. The invention also provides a method by which various PHA synthases can be tested for their specificity with respect to different monomer substrates.

The potential uses and applications of PHAs produced by PHA monomer synthases and PHA synthases includes both medical and industrial applications. Medical applications of PHAs include surgical pins, sutures, staples, swabs, wound dressings, blood vessel replacements, bone replacements and plates, stimulation of bone growth by piezoelectric properties, and biodegradable carrier for long-term dosage of pharmaceuticals. Industrial applications of PHAs include disposable items such as baby diapers, packaging containers, bottles, wrappings, bags, and films, and biodegradable carriers for long-term dosage of herbicides, fungicides, insecticides, or fertilizers.

In animals, the biosynthesis of fatty acids *de novo* from malonyl-CoA is catalyzed by FAS. For example, the rat FAS is a homodimer with a subunit structure consisting of 2505 amino acid residues having a molecular weight of 272,340 Da. Each subunit consists of seven catalytic activities in separate physical domains (Amy et al., *Proc. Natl. Acad. Sci. USA*, 86, 3114 (1989)). The physical location of six of the catalytic activities, ketoacyl synthase (KS), malonyl/acetyltransferase (M/AT), enoyl reductase (ER), ketoreductase (KR), acyl carrier protein (ACP), and thioesterase (TE), has been established by (1) the identification of the various active site residues within the overall amino acid sequence by isolation of

catalytically active fragments from limited proteolytic digests of the whole FAS, (2) the identification of regions within the FAS that exhibit sequence similarity with various monofunctional proteins, (3) expression of DNA encoding an amino acid sequence with catalytic 5 activity to produce recombinant proteins, and (4) the identification of DNA that does not encode catalytic activity, i.e., DNA encoding a linker region. (Smith et al., Proc. Natl. Acad. Sci. USA, 73, 1184 (1976); Tsukamoto et al., J. Biol. Chem., 263, 16225 (1988); Rangan et al., J. Biol. Chem., 266, 19180 (1991)).

10 The seventh catalytic activity, dehydrase (DH), was identified as physically residing between AT and ER by an amino acid comparison of FAS with the amino acid sequences encoded by the three open reading frames of the *eryA* polyketide synthase (PKS) gene cluster of *Saccharopolyspora erythraea*. The three polypeptides that 15 comprise this PKS are constructed from "modules" which resemble animal FAS, both in terms of their amino acid sequence and in the ordering of the constituent domains (Donadio et al., Gene, 111, 51 (1992); Benh et al., Eur. J. Biochem., 204, 39 (1992)).

One embodiment of the invention employs a FAS in 20 which the DH is inactivated (FAS DH⁻). The FAS DH⁻ employed in this embodiment of the invention is preferably a eukaryotic FAS DH⁻ and, more preferably, a mammalian FAS DH⁻. The most preferred embodiment of the invention is a FAS where the active site in the DH has been inactivated by mutation. For example, Joshi et al. (J. Biol. Chem., 268, 22508 (1993)) changed the His⁸⁷⁸ residue in the rat FAS to an alanine residue by site directed mutagenesis. *In vitro* studies showed that a FAS with this change (ratFAS206) produced 3-hydroxybutyrylCoA as a premature termination product from acetyl-CoA, malonyl-CoA and NADPH.

As shown below, a FAS DH- effectively replaces the β -ketothiolase and acetoacetyl-CoA reductase activities of the natural pathway by producing D(-)-3-hydroxybutyrate as a premature termination product, rather than the usual 16-carbon product,
5 palmitic acid. This premature termination product can then be incorporated into PHB by a PHB synthase (See Example 2).

Another embodiment of the invention employs a recombinant *Streptomyces spp.* PKS to produce a variety of β -hydroxyCoA esters that can serve as monomers for a PHA synthase.
10 One example of a DNA encoding a Type I PKS is the *eryA* gene cluster, which governs the synthesis of erythromycin aglycone deoxyerythronolide B (DEB). The gene cluster encodes six repeated units, termed modules or synthase units (SUs). Each module or SU, which comprises a series of putative FAS-like activities, is responsible
15 for one of the six elongation cycles required for DEB formation. Thus, the processive synthesis of asymmetric acyl chains found in complex polyketides is accomplished through the use of a programmed protein template, where the nature of the chemical reactions occurring at each point is determined by the specificities in each SU.

20 Two other Type I PKS are encoded by the *tyl* (tylosin) (Figure 4) and *met* (methymycin) (Figure 5) gene clusters. The macrolide multifunctional synthases encoded by *tyl* and *met* provide a greater degree of metabolic diversity than that found in the *eryA* gene cluster. The PKSs encoded by the *eryA* gene cluster only catalyze
25 chain elongation with methylmalonylCoA, as opposed to *tyl* and *met* PKSs, which catalyze chain elongation with malonylCoA, methylmalonylCoA and ethylmalonylCoA. Specifically, the *tyl* PKS includes two malonylCoA extender units and one ethylmalonylCoA extender unit, and the *met* PKS includes one malonylCoA extender
30 unit. Thus, a preferred embodiment of the invention includes, but is

not limited to, replacing catalytic activities encoded in *met* PKS open reading frame 1 (ORF1) to provide a DNA encoding a protein that possesses the required keto group processing capacity and short chain acylCoA ester starter and extender unit specificity necessary to provide
5 a saturated β -hydroxyhexanoylCoA or unsaturated β -hydroxyhexenoylCoA monomer.

In order to manipulate the catalytic specificities within each module, DNA encoding a catalytic activity must remain undisturbed. To identify the amino acid sequences between the
10 amino acid sequences with catalytic activity, the "linker regions," amino acid sequences of related modules, preferably those encoded by more than one gene cluster, are compared. Linker regions are amino acid sequences which are less well conserved than amino acid sequences with catalytic activity. Witkowski et al., Eur. J. Biochem.,
15 198, 571 (1991).

In an alternative embodiment of the invention, to provide a DNA encoding a Type I PKS module with a TE and lacking a functional DH, a DNA encoding a module F, containing KS, MT, KR, ACP, and TE catalytic activities, is introduced at the 3' end of a
20 DNA encoding a first module (Figure 6). Module F introduces the final (R)-3-hydroxyl acyl group at the final step of PHA monomer synthesis, as a result of the presence of a TE domain. DNA encoding a module F is not present in the *eryA* PKS gene cluster (Donadio et al., *supra*, 1991).

25 A DNA encoding a recombinant monomer synthase is inserted into an expression vector. The expression vector employed varies depending on the host cell to be transformed with the expression vector. That is, vectors are employed with transcription, translation and/or post-translational signals, such as targeting signals,
30 necessary for efficient expression of the genes in various host cells

into which the vectors are introduced. Such vectors are constructed and transformed into host cells by methods well known in the art. See Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor (1989). Preferred host cells for the vectors of the invention include insect, bacterial, and plant cells. Preferred insect cells include *Spodoptera frugiperda* cells such as *Sf21*, and *Trichoplusia ni* cells. Preferred bacterial cells include *Escherichia coli*, *Streptomyces* and *Pseudomonas*. Preferred plant cells include monocot and dicot cells, such as maize, rice, wheat, tobacco, legumes, 5 carrot, squash, canola, soybean, potato, and the like.

Moreover, the appropriate subcellular compartment in which to locate the enzyme in eukaryotic cells must be considered when constructing eukaryotic expression vectors. Two factors are important: the site of production of the acetyl-CoA substrate, and the 10 available space for storage of the PHA polymer. To direct the enzyme to a particular subcellular location, targeting sequences may be added to the sequences encoding the recombinant molecules.

The baculovirus system is particularly amenable to the introduction of DNA encoding a recombinant FAS or a PKS monomer synthase because an increasing variety of transfer plasmids 15 are becoming available which can accommodate a large insert, and the virus can be propagated to high titers. Moreover, insect cells are adapted readily to suspension culture, facilitating relatively large scale recombinant protein production. Further, recombinant proteins tend 20 to be produced exclusively as soluble proteins in insect cells, thus, obviating the need for refolding, a task that might be particularly daunting in the case of a large multifunctional protein. The *Sf21*/baculovirus system has routinely expressed milligram quantities 25 of catalytically active recombinant fatty acid synthase. Finally, the 30 baculovirus/insect cell system provides the ability to construct and

analyze different synthase proteins for the ability to polymerize monomers into unique biodegradable polymers.

A further embodiment of the invention is the introduction of at least one DNA encoding a PHA synthase and a 5 DNA encoding a PHA monomer synthase into a host cell. Such synthases include, but are not limited to, *A. eutrophus* 3-hydroxy, 4-hydroxy, and 5-hydroxy alkanoate synthases, *Rhodococcus ruber* C₃-C₅ hydroxyalkanoate synthases, *Pseudomonas oleororans* C₆-C₁₄ hydroxyalkanoate synthases, *P. putida* C₆-C₁₄ hydroxyalkanoate 10 synthases, *P. aeruginosa* C₅-C₁₀ hydroxyalkanoate synthases, *P. resinovorans* C₄-C₁₀ hydroxyalkanoate synthases, *Rhodospirillum rubrum* C₄-C₇ hydroxyalkanoate synthases, *R. gelatinorus* C₄-C₇, *Thiocapsa pfennigii* C₄-C₈ hydroxyalkanoate synthases, and *Bacillus megaterium* C₄-C₅ hydroxyalkanoate synthases.

15 The introduction of DNA(s) encoding more than one PHA synthase may be necessary to produce a particular PHA polymer due to the specificities exhibited by different PHA synthases. As multifunctional proteins are altered to produce unusual monomeric structures, synthase specificity may be problematic for particular substrates. Although the *A. eutrophus* PHB synthase utilizes only C₄ 20 and C₅ compounds as substrates, it appears to be a good prototype synthase for initial studies since it is known to be capable of producing copolymers of 3-hydroxybutyrate and 4-hydroxybutyrate (Kunioka et al., *Macromolecules*, 22, 694 (1989)) as well as copolymers of 3- 25 hydroxyvalerate, 3-hydroxybutyrate, and 5-hydroxyvalerate (Doi et al., *Macromolecules*, 19, 2860 (1986)). Other synthases, especially those of *Pseudomonas aeruginosa* (Timm et al., *Eur. J. Biochem.*, 209, 15 (1992)) and *Rhodococcus ruber* (Pieper et al., *FEMS Microbiol. Lett.*, 96, 73 (1992)), can also be employed in the practice of the invention. 30 Synthase specificity may be alterable through molecular biological methods.

In yet another embodiment of the invention, a DNA encoding a FAS and a PHA synthase can be introduced into a single expression vector, obviating the need to introduce the genes into a host cell individually.

5 A further embodiment of the invention is the generation of a DNA encoding a recombinant multifunctional protein, which comprises a FAS, of either eukaryotic or prokaryotic origin, and a PKS module F. Module F will carry out the final chain extension to include two additional carbons and the reduction of the β -keto group,
10 which results in a (R)-3-hydroxy acyl CoA moiety.

To produce this recombinant protein, DNA encoding the FAS TE is replaced with a DNA encoding a linker region which is normally found in the ACP-KS interdomain region of bimodular ORFs. DNA encoding a module F is then inserted 3' to the DNA
15 encoding the linker region. Different linker regions, such as those described below, which vary in length and amino acid composition, can be tested to determine which linker most efficiently mediates or allows the required transfer of the nascent saturated fatty acid intermediate to module F for the final chain elongation and keto
20 reduction steps. The resulting DNA encoding the protein can then be tested for expression of long chain β -hydroxy fatty acids in insect cells, such as *Sf21* cells, or *Streptomyces*, or *Pseudomonas*. The expected 3-hydroxy C-18 fatty acid can serve as a potential substrate for PHA synthases which are able to accept long chain alkyl groups. A
25 preferred embodiment of the invention is a FAS that has a chain length specificity between 4-22 carbons.

Examples of linker regions that can be employed in this embodiment of the invention include, but are not limited to, the ACP-KS linker regions encoded by the *tyl* ORFI (ACP₁-KS₂; ACP₂-

KS₃), and ORF3 (ACP₅-KS₆), and *eryA* ORFI (ACP₁-KS₁; ACP₂-KS₂), ORF2 (ACP₃-KS₄) and ORF3 (ACP₅-KS₆).

This approach can also be used to produce shorter chain fatty acid groups by limiting the ability of the FAS unit to generate 5 long chain fatty acids. Mutagenesis of DNA encoding various FAS catalytic activities, starting with the KS, may result in the synthesis of short chain (R)-3-hydroxy fatty acids.

The PHA polymers are then recovered from the biomass. Large scale solvent extraction can be used, but is expensive. An 10 alternative method involving heat shock with subsequent enzymatic and detergent digestive processes is also available (Byron, Trends Biotechnical, 5, 246 (1987); Holmes, In: Developments in Crystalline Polymers, D.C. Bassett (ed), pp. 1-65 (1988)). PHB and other PHAs are readily extracted from microorganisms by chlorinated hydrocarbons. 15 Refluxing with chloroform has been extensively used; the resulting solution is filtered to remove debris and concentrated, and the polymer is precipitated with methanol or ethanol, leaving low-molecular-weight lipids in solution. Longer-side-chain PHAs show a less restricted solubility than PHB and are, for example, soluble in 20 acetone. Other strategies adopted include the use of ethylene carbonate and propylene carbonate as disclosed by Lafferty et al. (Chem. Rundschau, 30, 14 (1977)) to extract PHB from biomass. Scandola et al., (Int. J. Biol. Microbiol., 10, 373 (1988)) reported that 1 M 25 HCl-chloroform extraction of *Rhizobium meliloti* yielded PHB of M_w = 6 x 10⁴ compared with 1.4 x 10⁶ when acetone was used.

Methods are well known in the art for the determination of the PHB or PHA content of microorganisms, the composition of PHAs, and the distribution of the monomer units in the polymer. Gas chromatography and high-pressure liquid chromatography are 30 widely used for quantitative PHB analysis. See Anderson et al. (Microbiol. Rev., 54, 450 (1990) for a review of such methods. NMR

techniques can also be used to determine polymer composition, and the distribution of monomer units.

The invention has been described with reference to various specific and preferred embodiments and will be further 5 described by reference to the following detailed examples. It is understood however, that there are many extensive variations, and modifications on the basic theme of the present invention beyond that shown in the examples and description, which are within the spirit and scope of the present invention.

10

I. Experimental Procedures

Materials and Methods

15 **Materials.** Sodium R-(-)-3-hydroxybutyrate, coenzyme-A, ethylchloroformate, pyridine and diethyl ether were purchased from Sigma Chemical Co. Amberlite IR-120 was purchased from Mallinckrodt Inc. 6-O-(N-Heptylcarbamoyl)methyl α-D-glucopyranoside (Hecameg) was obtained from Vinatec (Villejuif, France). Two-piece spectrophotometer cells with pathlengths of 0.1 (#20/0-Q-1) and 0.01 cm (#20/0-Q-0.1) were obtained from Starna Cells 20 Inc., (Atascadero, CA). Rabbit anti-*A. eutrophus* PHA synthase antibody was a gracious gift from Dr. F. Srienc and S. Stoup (Biological Process Technology Institute, University of Minnesota). Sf21 cells and 25 *T. ni* cells were kindly provided by Greg Franzen (R&D Systems, Minneapolis, MN) and Stephen Harsch (Department of Veterinary Pathobiology, University of Minnesota), respectively.

Plasmid pFAS206 and a recombinant baculoviral clone encoding FAS206 (Joshi et al., *J. Biol. Chem.*, **268**, 22508 (1993)) were 30 generous gifts of A. Joshi and S. Smith. Plasmid pAet41 (Peoples et al.,

J. Biol. Chem., **264**, 15298, (1989)), the source of the *A. eutrophus* PHB synthase, was obtained from A. Sinskey. Baculovirus transfer vector, pBacPAK9, and linearized baculoviral DNA, were obtained from Clontech Inc. (Palo Alto, CA). Restriction enzymes, T4 DNA ligase, *E. coli* DH5 α competent cells, molecular weight standards, lipofectin reagent, Grace's insect cell medium, fetal bovine serum (FBS), and antibiotic/antimycotic reagent were obtained from GIBCO-BRL (Grand Island, NY). Tissue culture dishes were obtained from Corning Inc. Spinner flasks were obtained from Bellco Glass Inc. Seaplaque agarose GTG was obtained from FMC Bioproducts Inc.

Methods.

Preparation of R-3HBCoA. R-(-)-3 HBCoA was prepared by the mixed anhydride method described by Haywood et al., *FEMS Microbiol. Lett.*, **57**, 1 (1989). 60 mg (0.58 mmol) of R-(-)-3 hydroxybutyric acid was freeze dried and added to a solution of 72 mg of pyridine in 10 ml diethyl ether at 0°C. Ethylchloroformate (100 mg) was added, and the mixture was allowed to stand at 4°C for 60 minutes. Insoluble pyridine hydrochloride was removed by centrifugation. The resulting anhydride was added, dropwise with mixing, to a solution of 100 mg coenzyme-A (0.13 mmol) in 4 ml 0.2 M potassium bicarbonate, pH 8.0 at 0°C. The reaction was monitored by the nitroprusside test of Stadtman, *Meth. Enzymol.*, **3**, 931 (1957), to ensure sufficient anhydride was added to esterify all the coenzyme-A. The concentration of R-3-HBCoA was determined by measuring the absorbance at 260 nm ($e = 16.8 \text{ mM}^{-1} \text{ cm}^{-1}$; 18).

Construction of pBP-phbC. The *phbC* gene (approximately 1.8 kb) was excised from pAet41 (Peoples et al., *J. Biol. Chem.*, **264**, 15293 (1989)) by digestion with *Bst*BI and *Stu*I, purified as described by Williams et al. (*Gene*, **109**, 445 (1991)), and ligated to pBacPAK9

digested with *Bst*BI and *Stu*I. This resulted in pBP-phbC, the baculovirus transfer vector used in formation of recombinant baculovirus particles carrying *phbC*.

Large scale expression of PHA synthase. A 1 L culture of *T. ni* cells (1.2×10^6 cells/ml) in logarithmic growth was infected by the addition of 50 ml recombinant viral stock solution (2.5×10^8 pfu/ml) resulting in a multiplicity of infection (MOI) of 10. This infected culture was split between two Bellco spinners (350 ml/500 ml spinner, 700 ml/1 L spinner) to facilitate oxygenation of the culture. These cultures were incubated at 28°C and stirred at 60 rpm for 60 hours. Infected cells were harvested by centrifugation at $1000 \times g$ for 10 minutes at 4°C. Cells were flash-frozen in liquid N₂ and stored in 4 equal aliquots, at -80°C until purification.

Insect cell maintenance and recombinant baculovirus formation. *Sf21* cells were maintained at 26-28°C in Grace's insect cell medium supplemented with 10% FBS, 1.0% pluronic F68, and 1.0% antibiotic/antimycotic (GIBCO-BRL). Cells were typically maintained in suspension at $0.2 - 2.0 \times 10^6$ /ml in 60 ml total culture volume in 100 ml spinner flasks at 55-65 rpm. Cell viability during the culture period was typically 95-100%. The procedures for use of the transfer vector and baculovirus were essentially those described by the manufacturer (Clontech, Inc.). Purified pBP-phbC and linearized baculovirus DNA were used for cotransfection of *Sf21* cells using the liposome mediated method (Felgner et al., Proc. Natl. Acad. Sci. USA, 84, 7413 (1987)) utilizing Lipofectin (GIBCO-BRL). Four days later cotransfection supernatants were utilized for plaque purification. Recombinant viral clones were purified from plaque assay plates containing 1.5% Seaplaque GTG after 5-7 days at 28°C. Recombinant viral clone stocks were then amplified in T25-flask cultures (4 ml, 3×10^6 /ml on day 0) for 4 days; infected cells were determined by their morphology and size and then screened by SDS/PAGE using 10%

polyacrylamide gels (Laemmli, *Nature*, **227**, 680 (1970)) for production of PHA synthase.

Purification of PHA synthase from BTI-TN-5B1-4 *T. ni* cells. Purification of PHA synthase was performed according to the method of Gerngross et al., *Biochemistry*, **33**, 9311 (1994) with the following alterations. One aliquot (110 mg protein) of frozen cells was thawed on ice and resuspended in 10 mM KPi (pH 7.2), 5% glycerol, and 0.05% Hecameg (Buffer A) containing the following protease inhibitors at the indicated final concentrations: benzamidine (2mM), phenylmethylsulfonyl fluoride (PMSF, 0.4 mM), pepstatin (2 mg/ml), leupeptin (2.5 mg/ml), and Na-p-tosyl-l-lysine chloromethyl ketone (TLCK, 2 mM). EDTA was omitted at this stage due to its incompatibility with hydroxylapatite (HA). This mixture was homogenized with three series of 10 strokes each in two Thomas 15 homogenizers while partially submerged in an ice bath and then sonicated for 2 minutes in a Branson Sonifier 250 at 30% cycle, 30% power while on ice. All subsequent procedures were carried out at 15 4°C.

The lysate was immediately centrifuged at 100000 x g in a Beckman 50.2Ti rotor for 80 minutes, and the resulting supernatant (10.5 ml, 47 mg) was immediately filtered through a 0.45 mm Uniflow filter (Schleicher and Schuell Inc., Keene, N.H.) to remove any remaining insoluble matter. Aliquots of the soluble fraction (1.5 ml, 7 mg) were loaded onto a 5 ml BioRad Econo-Pac HTP column that had been equilibrated with Buffer A (+ protease inhibitor mix) attached to a BioRad Econo-system, and the column was washed with 30 ml Buffer A. All chromatographic steps were carried out at a flow rate of 0.8 ml/minute. PHA synthase was eluted from the HA column with a 32 x 32 ml linear gradient from 10 to 300 mM KPi.

30 Fraction collection tubes were prepared by addition of 30 ml of 100 mM EDTA to provide a metalloprotease inhibitor at 1 mM

immediately after HA chromatography. PHA synthase was eluted in a broad peak between 110-180 mM KPi. Fractions (3 ml) containing significant PHA synthase activity were pooled and stored at 0°C until the entire soluble fraction had been run through the chromatographic process. Pooled fractions then were concentrated at 4°C by use of a Centriprep-30 concentrator (Amicon) to 3.8 mg/ml. Aliquots (0.5 ml) were either flash-frozen and stored in liquid N₂ or glycerol was added to a final concentration of 50% and samples (1.9 mg/ml) were stored at -20°C.

Western analysis. Samples of *T. ni* cells were fractionated by SDS-PAGE on 10% polyacrylamide gels, and the proteins then were transferred to 0.2 mm nitrocellulose membranes using a BioRad Transblot SD Semi-Dry electrophoretic transfer cell according to the manufacturer. Proteins were transferred for 1 hour at 15 V. The membrane was rinsed with doubly distilled H₂O, dried, and treated with phosphate-buffered saline (PBS) containing 0.05% Tween-20 (PBS-Tween) and 3% nonfat dry milk to block non-specific binding sites. Primary antibody (rabbit anti-PHA synthase) was applied in fresh blocking solution and incubated at 25°C for 2 hours. Membranes were then washed four times for 10 minutes with PBS-Tween followed by the addition of horseradish peroxidase-conjugated goat-anti-rabbit antibody (Boehringer-Mannheim) diluted 10,000X in fresh blocking solution and incubated at 25°C for 1 hour. Membranes were washed finally in three changes (10 minutes) of PBS, and the immobilized peroxidase label was detected using the chemiluminescent LumiGLO substrate kit (Kirkegaard and Perry, Galthersburg, MD) and X-ray film.

N-terminal analysis. Approximately 10 mg of purified PHA synthase was run on a 10% SDS-polyacrylamide gel, transferred to PVDF (Immobilon-PSQ, Millipore Corporation, Bedford, MA), stained with Amido Black, and sequenced on a 494 Procise Protein

Sequencer (Perkin-Elmer, Applied Biosystems Division, Foster City, California).

Double-infection protocol. Four 100 ml spinner flasks were each inoculated with 8×10^7 cells in 50 ml of fresh insect medium. To flask 1, an additional 20 ml of fresh insect medium was added (uninfected control); to flask 2, 10 ml BacPAK6::*phbC* viral stock (1×10^8 pfu/ml) and 10 ml fresh insect medium were added; to flask 3, 10 ml BacPAK6::FAS206 viral stock (1×10^8 pfu/ml) and 10 ml fresh insect medium were added; and to flask 4, 10 ml BacPAK6::*phbC* viral stock (1×10^8 pfu/ml) and 10 ml BacPAK6::FAS206 viral stock (1×10^8 pfu/ml) were added. These viral infections were carried out at a multiplicity of infection of approximately 10. Cultures were maintained under normal growth conditions and 15 ml samples were removed at 24, 48, and 72 hour time points. Cells were collected by gentle centrifugation at $1000 \times g$ for 5 minutes, the medium was discarded, and the cells were immediately stored at -70°C .

PHA synthase assays. Coenzyme A released by PHA synthase in the process of polymerization was monitored precisely as described by Gerngross et al. (*supra*) using 5,5'-dithiobis (2-nitrobenzoic acid, DTNB) (Ellman, *Arch. Biochem. Biophys.*, **82**, 70 (1959)).

The presence of HBCoA was monitored spectrophotometrically. Assays were performed at 25°C in a Hewlett Packard 8452A diode array spectrophotometer equipped with a water jacketed cell holder. Two-piece Starna Spectrosil spectrophotometer cells with pathlengths of 0.1 and 0.01 cm were employed to avoid errors arising from the compression of the absorbance scale at higher values. Absorbance was monitored at 232 nm, and $E_{232\text{nm}}$ of 4.5×10^3 M $^{-1}$ cm $^{-1}$ was used in calculations. One unit (U) of enzyme is the amount required to hydrolyze 1 mmol of substrate minute $^{-1}$. Buffer

(0.15 M KPi, pH 7.2) and substrate were equilibrated to 25°C and then combined in an Eppendorf tube also at 25°C. Enzyme was added and mixed once in the pipet tip used to transfer the entire mixture to the spectrophotometer cell. The two piece cell was immediately 5 assembled, placed in the spectrophotometer with the cell holder (type CH) adapted for the standard 10 mm path length cell holder of the spectrophotometer. Manipulations of sample, from mixing to initiation of monitoring, took only 10-15 seconds. Absorbance was continually monitored for up to 10 minutes. Calibration of reactions 10 was against a solution of buffer and enzyme (no substrate) which lead to absorbance values that represented substrate only.

PHB assay. PHB was assayed from *Sf21* cell samples according to the propanolysis method of Riis et al., *J. Chromo.*, **445**, 285 (1988). Cell pellets were thawed on ice, resuspended in 1 ml cold 15 ddH₂O and transferred to 5 ml screwtop test tubes with teflon seals. 2 ml ddH₂O was added, the cells were washed and centrifuged and then 3 ml of acetone were added and the cells washed and centrifuged. The samples were then dessicated by placing them in a 94°C oven for 12 hours. The following day 0.5 ml of 1,2-dichloroethane, 0.5 ml 20 acidified propanol (20 ml HCl, 80 ml 1-propanol) and 50 ml benzoic acid standard were added and the sealed tubes were heated to 100°C in a boiling water bath for 2 hours with periodic vortexing. The tubes were cooled to room temperature and the organic phase was used for gas-chromatographic (GC) analysis using a Hewlett Packard 5890A gas- 25 chromatograph equipped with a Hewlett Packard 7673A automatic injector and a fused silica capillary column, DB-WAX 30W of 30 meter length. Positive samples were further subjected to GC-mass spectrometric (MS) analysis for the presence of propylhydroxybutyrate using a Kratos MS25 GC/MS. The following parameters were used: 30 source temperature, 210°C; voltage, 70eV; and accelerating voltage, 4 KeV.

Catalytic activities.

Ketoacyl synthase (KS) activity was assessed radiochemically by the condensation-¹⁴CO₂ exchange reaction (Smith et al., PNAS USA, **73**, 1184 (1976)).

5 Transferase (AT) activity was assayed, using malonyl-CoA as donor and pantetheine as acceptor, by determining spectrophotometrically the free CoA released in a coupled ATP citrate-lyase-malate dehydrogenase reaction (see, Rangen et al., J. Biol. Chem., **266**, 19180 (1991)).

10 Ketoreductase (KR) was assayed spectrophotometrically at 340 nm: assay systems contained 0.1 M potassium phosphate buffer (pH 7), 0.15 mM NADPH, enzyme and either 10 mM *trans*-1-decalone or 0.1 mM acetoacetyl-CoA substrate.

15 Dehydrase (DH) activity was assayed spectrophotometrically at 270 nm using S-DL-β-hydroxybutyryl N-acetylcysteamine as substrate (Kumar et al., J. Biol. Chem., **245**, 4732 (1970)).

20 Enoyl reductase (ER) activity was assayed spectrophotometrically at 340 nm essentially as described by Strometal (J. Biol. Chem., **254**, 8159 (1979)); the assay system contained 0.1 M potassium phosphate buffer (pH 7), 0.15 mM NADPH, 0.375 mM crotonoyl-CoA, 20 μM CoA and enzyme.

25 Thioesterase (TE) activity was assessed radiochemically by extracting and assaying the [¹⁴C]palmitic acid formed from [1-¹⁴C]palmitoyl-CoA during a 3 minute incubation Smith, Meth. Enzymol., **71C**, 181 (1981); the assay was in a final volume of 0.1 ml, 25 mM potassium phosphate buffer (pH 8), 20 μM [1-¹⁴C]palmitoyl-CoA (20 nCi) and enzyme.

30 Assay of overall fatty acid synthase activity was performed spectrophotometrically as described previously by Smith et al. (Meth.

Enzymol., 35, 65 (1975)). All enzyme activities were assayed at 37°C except the transferase, which was assayed at 20°C. Activity units indicate nmol of substrate consumed/minute. All assays were conducted, at a minimum, at two different protein concentrations
5 with the appropriate enzyme and substrate blanks included.

Example 1

Expression of *A. eutrophus* PHA synthase using a baculovirus system.

Recent work has shown that PHA synthase from *A. eutrophus* can be overexpressed in *E. coli*, in the absence of 3-ketothiolase and acetoacetyl-CoA reductase (Gerngross et al. *supra*) and can be expressed in plants (See Poirier et al., Biotech, 13, 142 (1995) for a review). Isolation of the soluble form of PHA synthase provides opportunities to examine the mechanistic details of the priming and initiation reactions. Because the baculovirus system has been successful for the expression of a number of prokaryotic genes as soluble proteins, and insect cells, unlike bacterial expression systems, carry out a wide array of posttranslational modifications, the baculovirus expression system appeared ideal for the expression of large quantities of soluble PHA synthase, a protein that must be modified by phosphopantetheine in order to be catalytically active (Gerngross et al., *supra*).

Purification of PHA synthase. The purification procedure employed for PHA synthase is a modification of Gerngross et al. (*supra*) involving the elimination of the second liquid chromatographic step and inclusion of a protease-inhibitor cocktail in all buffers. All steps were carried out on ice or at 4°C except where noted. Frozen cells were thawed on ice in 10 ml of Buffer A (10 mM KPi, pH 7.2, 0.5% glycerol, and 0.05% Hecameg) and then immediately homogenized prior to centrifugation and HA chromatography.

The results of these efforts are summarized in Table 1 and Figure 7. A prominent band at 64 kDa is visible in total, soluble, and HA eluate protein samples fractionated by SDS/PAGE (lanes 4, 5, and 6 of Figure 7, respectively). The initial specific activity of the isolated PHA synthase was 20-fold higher than previous attempts at expression and purification of this polypeptide. Approximately 1000 units of PHB synthase have been purified, based on calculations from the direct spectrophotometric assay detailed below, with an overall recovery of activity of 70%. The large proportion of synthase present in the membrane fraction, and the fact that over 90% of the initial activity was found in the soluble fraction, suggests either that the synthase in the membrane fraction is in an inactive form or that the direct assay is not applicable to the initial, 12 U/mg, crude extract.

15 Table 1: Purification of PHA Synthase

sample	total units	vol (mL)	protein (mg)	specific (mg/ml)	activity	recovery
total protein	1430	11.5	113	9.8	12.7	100
soluble protein	1340	10.5	47	4.5	28.6	93
pooled HA fractions	1020	7.9	30	3.8	34.2	71
-----	-----	-----	-----	-----	-----	-----

20 N-terminal sequencing of the 64 kDa protein confirmed its identity as PHA synthase (Figure 8). Two prominent N-termini, at amino acid residue 7 (alanine) and residue 10 (serine) were obtained in a 3:2 ratio. This heterogeneous N-terminus presumably is the result of aminopeptidase activity. Western analysis using a rabbit-

anti-PHA synthase antibody corroborated the results of the sequencing and indicated the presence of at least three bands that resulted from proteolysis of PHA synthase (Figure 7B, Lanes 4-6). The antibody was specific for PHA synthase since neither *T. ni* nor baculoviral proteins 5 showed reactivity (Figure 7B, Lanes 2 and 3). N-terminal protein sequencing (Figure 8) showed directly that the 44 kDa (band b) and 32 kDa (band d) proteins were derived from PHA synthase (fragments beginning at A181/N185 and at G387, respectively). The 35-40 kDa (band c) protein gave low sequencing yields and may contain a 10 blocked N-terminus. Inspection of Figure 7B suggests that most degradation occurs following cell disruption since the total protein sample for this gel (lane 4) was prepared by boiling intact cells directly in SDS sample buffer while the HA sample (lane 6) went through the purification procedure described above.

15 **Assay of Synthase Activity.** Due to the significant level of expression obtained using the baculovirus system, the synthase activity could be assayed spectrophotometrically by monitoring hydrolysis of the thioester bond at 232 nm, the wavelength at which there is a maximum decrease in absorbance upon hydrolysis. The 20 difference between substrate (HBCoA) and product (CoA) at this wavelength is shown in Figure 9. Absorbance of HBCoA and CoA at 232 nm occurs at a trough between two well separated peaks. Assays were carried out at pH 7.2 for comparative analysis with previous studies (Gerngross et al., *supra*). Substrate (R-(-)-3-HBCoA) substrate for 25 these studies was prepared using the mixed anhydride method (Haywood et al., *supra*), and its concentration was determined by measuring A_{260} . The short pathlength cells (0.1 cm and 0.01 cm) allowed use of relatively high reaction concentrations while conserving substrate and enzyme. Assay results showed an initial lag 30 period of 60 seconds prior to the linear decrease in A_{232} , and velocities were determined from the slope of these linear regions of

the assay curves. The length of the lag period was variable and was inversely related to enzyme concentration. These data are consistent with those using PHA synthase purified from *E. coli* (Gerngross et al., *supra*).

5 Figures 10 and 11 show the V versus S and 1/V versus 1/S plots, respectively. The double reciprocal plot was concave upward which is similar to results obtained from studies of the granular PHA synthase from *Zooglea ramigera* (Fukui et al., Arch. Microbiol., **110**, 149 (1976)) and suggests a complex reaction mechanism. Examinations of velocity and specific activity as a function of enzyme concentration are shown in Figures 12 and 13. These results confirm that specific activity of the synthase depends upon enzyme concentration. The pH activity curve for *A. eutrophus* PHA synthase purified from *T. ni* cells is shown in Figure 14. The 10 curve shows a broad activity maximum centered around pH 8.5. This result agrees well with prior work on the *A. eutrophus* PHB synthase although it is significantly different than results obtained for the PHB synthase from *Z. ramigera* for which the optimum was determined to be pH 7.0.

15

20 The effect of varying enzyme concentration in the presence of a fixed amount of substrate revealed an intriguing trend (Figure 15). From these data it appears that the extent of polymerization is dependent on the amount of enzyme included in the reaction mixture. This could be explained if there is a "terminal 25 length" limitation of the polymer, which, once reached, can not be extended any further. If this is the case, it would also suggest that termination of the polymerization reaction, the release of the synthase from the polymer, and/or reinitiation of polymerization by the newly released synthase are relatively slow events since no 30 evidence of these reactions are seen within the timecourse of these studies. The phenomenon observed in Figure 15 is not the result of

decay of the enzyme over the course of the assay since virtually identical results are obtained following a 10 minute preincubation of the synthase at 25°C.

It must also be noted that comparisons of the direct 5 spectrophotometric assays used here and the more common assay involving the use of Ellman's reagent, DTNB, (Ellman, *supra*) in the formation of thiolate of coenzyme-A showed that the values determined by the direct method were approximately 70% of the values determined using Ellman's reagent. This may be due to phase 10 separation occurring in the cuvettes as the relatively insoluble polymer is formed. In support of this notion, a faint haze or opalescence in the cuvette developed during the course of the reaction, particularly at higher substrate concentrations.

PHA synthase purified from insect cells appears to be 15 relatively stable. Examination of activity following storage, in liquid N₂ and at -20°C in the presence of 50% glycerol showed that approximately 50% of synthase activity remained after 7 weeks when stored in liquid N₂ and approximately 75% of synthase activity remained after 7 weeks when stored at -20°C in the presence of 50% 20 glycerol.

The expression of PHA synthase from *A. eutrophus* in a baculovirus expression system results in the synthase constituting approximately 50% of total protein 60 hours post-infection; however, approximately 50-75% of the synthase is observed in the membrane- 25 associated fraction. This elevated level of expression allowed purification of the soluble PHA synthase using a single chromatographic step on HA. The purity of this preparation is estimated to be approximately 90% (intact PHA synthase and 3 proteolysis products).

30 The initial specific activity of 12 U/mg was approximately 20-fold higher than the most successful previous

efforts at overexpression of *A. eutrophus* PHA synthase. The synthase reported here was isolated from a 250 ml culture with 70% recovery which represents an improvement of 500-fold (1000 U / 64 U x 8 L / 0.25 L) when compared to an 8 L *E. coli* culture with 40% recovery.

5 This high expression level should provide sufficient PHA synthase for extensive structural, functional, and mechanistic studies. Furthermore, it is clear that the baculovirus expression system is an attractive option for isolation of other PHA synthases from various sources.

10 PHA synthase produced in the baculovirus system was of sufficient potency to allow direct spectrophotometric analysis of the hydrolysis of the thioester bond of HBCoA at 232 nm. These assays revealed a lag period of approximately 60 seconds, the length of which was variable and inversely related to enzyme concentration. Such a
15 lag period presumably reflects a slow step in the reaction, perhaps correlating to dimerization of the enzyme, the priming, and/or initiation steps in formation of PHB. Size exclusion chromatographic examination of the PHB synthase native MW indicated two forms of the synthase. One form showed a MW of approximately 100-160 kDa
20 and the other showed a MW of approximately 50-80 kDa; these two forms likely represent the dimer and monomer of PHA synthase, respectively. Similar results have been reported previously in which two forms of approximately 60 and 130 kDa were observed.
25 Comparisons of the direct assay reported here and the indirect assay using DTNB revealed that the former resulted in values that were 70% of the values determined by the DTNB indirect assay. Although the reason for this difference has not been examined in detail, it is probable that the apparent phase separation that occurred upon PHB formation in the short pathlength cuvettes used, particularly with
30 high [HBCoA], results in this discrepancy.

Enzymatic analyses of the PHA synthase have found that the enzyme has a broad pH optimum centered at pH 8.5; however, the studies described herein have been performed at pH 7.2 to provide comparative values with the results of others. Moreover, the specific 5 activity of this enzyme is dependent upon enzyme concentration which confirms and extends earlier results (Gerngross et al., *supra*).

In studies intended to examine the dependence of activity upon enzyme concentration, it became apparent that the extent of the polymerization reaction is dependent on the amount of 10 enzyme included in the reaction mixture. Specifically, decreasing the amount of enzyme leads not only to decreased velocity of reaction but also to a decreased extent of condensation (Figure 15). One possible explanation is that the enzyme is thermally labile; however, identical assays in which the enzyme is preincubated at 25°C for 10 minutes 15 prior to initiation of the reaction had similar results. Another possibility is that a terminal-length of the polymer is reached precluding further condensations until the particular synthase molecule is released from the terminal-length polymer.

This work clearly demonstrates the value of the 20 baculovirus expression system for the production of *A. eutrophus* PHA synthase and for the potential application to studies of other PHA synthases. Furthermore, the high level of expression obtained using the baculoviral system should allow convenient analysis for substrate-specificity and structure-function studies of PHA synthases 25 from relatively crude insect cell extracts.

Example 2

Co-expression of rat FAS dehydrase mutant cDNA and PHB synthase gene in insect cells.

30 Expression of a rat FAS DH cDNA in *Sf9* cells has been reported previously (Rangan et al., *J. Biol. Chem.*, **266**, 19180 (1991);

Joshi et al., Biochem. J. **296**, 143 (1993)). Once activity of the *phbC* gene product had been established in insect cells (see Example 1), baculovirus clones containing the rat FAS DH⁺ cDNA and BacPAK6::*phbC* were employed in a double infection strategy to 5 determine if PHB would be produced in insect cells. It was not known if an intracellular pool of R(-)-3-hydroxybutyrate would be stable or available as a substrate for the PHB synthase. In order for the R(-)-3-hydroxybutyrylCoA to be available as a substrate, the R(-)-3-hydroxybutyrylCoA released from rat FAS DH⁺ protein must be 10 trapped by the PHB synthase and incorporated into a polymer at a rate faster than β -oxidation, which would regenerate acetylCoA. It was also not known if the stereochemical configuration of the 3-hydroxyl group, which must be in the R form, would be recognized as a substrate by PHB synthase. Fortunately, previous biochemical studies 15 on eukaryotic FASs indicated that the R form of 3-hydroxylbutyrylCoA would be generated (Wakil et al., J. Biol. Chem., **237**, 687 (1962)).

SDS-PAGE of protein samples from a time course of uninfected, single-infected, and dual-infected Sf21 cells was performed 20 (Figure 16). From these data, it is clear that the rat FAS DH mutant and PHB synthase polypeptides are efficiently co-expressed in Sf21 cells. However, co-expression results in ~50% reduced levels of both polypeptides compared to Sf21 cells that are producing the individual proteins. Western analysis using anti-rat FAS (Rangan et al., *supra*) 25 and anti-PHA synthase antibodies confirmed simultaneous production of the corresponding proteins.

To provide further evidence that PHB was being synthesized in insect cells, *T. ni* cells which had been infected with a baculovirus vector encoding rat FAS DH⁰ and/or a baculovirus vector 30 encoding PHA synthase were analyzed for the presence of granules.

Infected cells were fixed in paraformaldehyde and incubated with anti-PHA synthase antibodies (Williams et al., *Protein Exp. Purif.*, 7, 203 (1996)). Granules were observed only in doubly infected cells (Williams et al., *App. Environ. Micro.*, 62, 2540 (1996)).

5 Characterization of PHB production in insect cells. In order to determine if *de novo* synthesis of PHB was occurring in Sf21 cells that co-express the rat FAS DH mutant and PHB synthase, fractions of these samples were extracted, the extract subjected to propanolysis, and analyzed for the presence of propylhydroxybutyrate
10 by gas chromatography (Figure 17). A unique peak with a retention time that coincided with a propylhydroxybutyrate standard was detected only in the double infection samples at 48 and 72 hours, in contrast to the individually expressed gene products and uninfected controls, which were negative. These samples were analyzed further
15 by GC/MS to confirm the identity of the product. Figure 18 shows mass spectroscopy data corresponding to the material obtained from peak 10.1 in the gas chromatograph compared to an propylhydroxybutyrate standard. The results show that PHB synthesis is occurring only in Sf21 cells co-expressing the rat FAS DH mutant
20 cDNA and the *phbC* gene from *A. eutrophus*. Integration of the peak in the gas chromatograph corresponding to propylhydroxybutyrate revealed that approximately 1 mg of PHB was isolated from 1 liter culture of Sf21 cells (approximately 600 mg dry cell weight of Sf21 cells). Thus, the ratFAS206 protein effectively replaces the β -
25 ketothiolase and acetoacetyl-CoA reductase functions, resulting in the production of PHB by a novel pathway.

The approach described here provides a new strategy to combine metabolic pathways that are normally engaged in primary anabolic functions for production of polyesters. The premature
30 termination of the normal fatty acid biosynthetic pathway to provide

suitably modified acylCoA monomers for use in PHA synthesis can be applied to both prokaryotic and eukaryotic expression since the formation of polymer will not be dependent on specialized feedstocks. Thus, once a recombinant PHA monomer synthase is introduced into 5 a prokaryotic or eukaryotic system, and co-expressed with the appropriate PHA synthase, novel biopolymer formation can occur.

Example 3

Cloning and Sequencing of the *vep* ORFI PKS Gene Cluster

10 The entire PKS cluster from *Streptomyces venezuelae* was cloned using a heterologous hybridization strategy. A 1.2 kb DNA fragment that hybridized strongly to a DNA encoding an *eryA* PKS β-ketoacyl synthase domain was cloned and used to generate a plasmid for gene disruption. This method generated a mutant strain blocked 15 in the synthesis of the antibiotic. A *S. venezuelae* genomic DNA library was generated. and used to clone a cosmid containing the complete methymycin aglycone PKS DNA. Fine-mapping analysis was performed to identify the order and sequence of catalytic domains along the multifunctional PKS (Figure 19). DNA sequence analysis of 20 the *vep* ORFI showed that the order of catalytic domains is KSQ/AT/ACP/KS/AT/KR/ACP/KS/AT/DH/KR/ACP. The complete DNA sequence, and corresponding amino acid sequence, of the *vep* ORFI is shown in Figure 23 (SEQ ID NO:1 and SEQ ID NO:2, respectively).

25 The sequence data indicated that the PKS gene cluster encodes a polyene of twelve carbons. The *vep* gene cluster contains 5 polyketide synthase modules, with a loading module at its 5' end and an ending domain at its 3' end. Each of the sequenced modules includes a keto-ACP (KS), an acyltransferase (AT), a dehydratase (DH),

a keto-reductase (KR), and an acyl carrier protein domain. The six acyltransferase domains in the cluster are responsible for the incorporation of six acetyl-CoA moieties into the product. The loading module contains a KSQ, an AT and an ACP domain. KSQ 5 refers to a domain that is homologous to a KS domain except that the active site cysteine (C) is replaced by glutamine (Q). There is no counterpart to the KSQ domain in the PKS clusters which have been characterized.

The ending domain (ED) is an enzyme which is 10 responsible for the attachment of the nascent polyketide chain onto another molecule. The amino acid sequence of ED resembles an enzyme, HetM, which is involved in *Anabaena* heterocyst formation. The homology between *vep* and HetM suggests that the polypeptide encoded by the *vep* gene cluster may synthesize a polyene-containing 15 composition which is present in the spore coat or cell wall of its natural host, *S. venezuelae*.

Example 4

To provide a recombinant monomer synthase that 20 generates a saturated β -hydroxyhexanoylCoA or unsaturated β -hydroxyhexanoylCoA monomer, the linear correspondence between the genetic organization of the Type I macrolide PKS and the catalytic domain organization in the multifunctional proteins is assessed (Donadio et al., *supra*, 1991; Katz et al., *Ann. Rev. Microbiol.*, **47**, 875 25 (1993)). First, a DNA encoding a TE is added to the 3' end of an ORF I of a Type I PKS, preferably the *met* ORF I (Figure 6) as recently described by Cortes et al. (*Science*, **268**, 1487 (1995) in the erythromycin system. To ensure that the DNA encoding the TE is completely active, DNA encoding a linker region separating a normal ACP-TE 30 region in a PKS, for example the one found in *met* PKS ORF5 (Figure

5), will be incorporated into the DNA. The resulting vector can be introduced into a host cell and the TE activity, rate of release of the CoA product, and identity of the fatty acid chain determined.

The acyl chain that is most likely to be released is the CoA ester, specifically the 3-hydroxy-4-methyl heptenoylCoA ester, since the fully elongated chain is presumably released in this form prior to macrolide cyclization. If the CoA form of the acyl chain is not observed, then a gene encoding a CoA ligase will be cloned and co-expressed in the host cell to catalyze formation of the desired 10 intermediate.

There is clear precedent for release of the predicted premature termination products from mutant strains of macrolide-producing *Streptomyces* that produce intermediates in macrolide synthesis (Huber et al., *Antimicrob. Agents Chemother.*, **34**, 1535 15 (1990); Kinoshita et al., *J. Chem. Soc. Chem. Comm.*, **14**, 943 (1988)). The structure of these intermediates is consistent with the linear organization of functional domains in macrolide PKSs, particularly those related to *eryA*, *tyl*, and *met*. Other known PKS gene clusters include, but are not limited to, the gene cluster encoding 6- 20 methysalicylic acid synthase (Beck et al., *Eur. J. Biochem.*, **192**, 487 (1990)), soraphen A (Schupp et al., *J. Bacteriol.*, **177**, 3673 (1995), and sterigmatocystin (Yu et al., *J. Bacteriol.*, **177**, 4792 (1995)).

Once the release of the 3-hydroxy-4-methyl heptenoylCoA ester is established, DNA encoding the extender unit AT in *met* module 1 is replaced to change the specificity from methylmalonylCoA to malonylCoA (Figures 4-6). This change eliminates methyl group branching in the β -hydroxy acyl chain. While comparison of known AT amino acid sequences shows high overall amino acid sequence conservation, distinct regions are readily 30 apparent where significant deletions or insertions have occurred. For

example, comparison of malonyl and methylmalonyl amino acid sequences reveals a 37 amino acid deletion in the central region of the malonyltransferase. Thus, to change the specificity of the methylmalonyl transferase to malonyl transferase, the *met* ORFI 5 DNA encoding the 37 amino acid sequence of MMT will be deleted, and the resulting gene will be tested in a host cell for production of the desmethyl species, 3-hydroxyheptenoylCoA. Alternatively, the DNA encoding the entire MMT can be replaced with a DNA encoding an intact MT to affect the desired chain construction.

10 After replacing MMT with MT, DNA encoding DH/ER will be introduced into DNA encoding *met* ORFI module 1. This modification results in a multifunctional protein that generates a methylene group at C-3 of the acyl chain (Figure 6). The DNA encoding DH/ER will be PCR amplified from the available *eryA* or *tyl* 15 PKS sequences, including the DNA encoding the required linker regions, employing a primer pair to conserved sequences 5' and 3' of the DNA encoding DH/ER. The PCR fragment will then be cloned into the *met* ORFI. The result is a DNA encoding a multifunctional protein (MT* DH/ER*TE*). This protein possesses the full 20 complement of keto group processing steps and results in the production of heptenoylCoA.

The DNA encoding dehydrase in *met* module 2 is then inactivated, using site-directed mutagenesis in a scheme similar to that used to generate the rat FAS DH described above (Joshi et al., L 25 *Biol. Chem.*, 268, 22508 (1993)). This preserves the required (R)-3-hydroxy group which serves as the substrate for PHA synthases and results in a (R)-3-hydroxyheptanoylCoA species.

The final domain replacement will involve the DNA 30 encoding the starter unit acyltransferase in *met* module 1 (Figure 5), to change the specificity from propionyl CoA to acetyl CoA. This shortens the (R)-3-hydroxy acyl chain from heptanoyl to hexanoyl.

The DNA encoding the catalytic domain will need to be generated based on a FAS or 6-methylsalicylic acid synthase model (Beck et al., *Eur. J. Biochem.*, **192**, 487 (1990)) or by using site-directed mutagenesis to alter the specificity of the resident *met* PKS propionyltransferase sequence. Limiting the initiator species to acetylCoA can result in the use of this starter unit by the monomer synthase. Previous work with macrolide synthases have shown that some are able to accept a wide range of starter unit carboxylic acids. This is particularly well documented for avermectin synthase, where over 60 new compounds have been produced by altering the starter unit substrate in precursor feeding studies (Dutton et al., *J. Antibiotics*, **44**, 357 (1991)).

Example 5

To provide a recombinant monomer synthase that synthesizes 3-hydroxyl-4-hexenoic acid, a precursor for polyhydroxyhexenoate, the DNA segment encoding the loading and the first module of the *vep* gene cluster was linked to the DNA segment encoding module 7 of the *tyl* gene cluster so as to yield a recombinant DNA molecule encoding a fusion polypeptide which has no amino acid differences relative to the corresponding amino acid sequence of the parent modules. The fusion polypeptide catalyzes the synthesis of 3-hydroxyl-4-hexenoic acid. The recombinant DNA molecule was introduced into SCP2, a *Streptomyces* vector, under the control of the *act* promoter (pDHS502, Figure 20). A polyhydroxyalkanoate polymerase gene, *phaC1* from *Pseudomonas oleavorans*, was then introduced downstream of the recombinant PKS cluster (pDHS505; Figures 22 and 23). The DNA segment encoding the polyhydroxyalkanoate polymerase is linked to the DNA segment encoding the recombinant PKS synthase so as to yield a fusion polypeptide which synthesizes polyhydroxyhexenoate in *Streptomyces*. Polyhydroxyhexenoate, a biodegradable thermoplastic,

is not naturally synthesized in *Streptomyces*, or as a major product in any other organism. Moreover, the unsaturated double bond in the side chain of polyhydroxyhexenoate may result in a polymer which has superior physical properties as a biodegradable thermoplastic over
5 the known polyhydroxyalkanoates.

The complete disclosure of all patents, patent documents and publications cited herein are incorporated herein by reference as if individually incorporated. The foregoing detailed description and examples have been given for clarity of understanding only. No
10 unnecessary limitations are to be understood therefrom. The invention is not limited to the exact details shown and described for variations obvious to one skilled in the art will be included within the invention defined by the claims.

WHAT IS CLAIMED IS:

1. A baculovirus expression cassette comprising a nucleic acid molecule encoding a polyhydroxyalkanoate synthase operably linked to a promoter functional in an insect cell.
2. The expression cassette of claim 1 wherein the source of the nucleic acid molecule is a bacterium.
3. The expression cassette of claim 2 wherein the bacterium is *Alcaligenes eutrophus*.
4. An expression cassette comprising a nucleic acid molecule encoding a polyhydroxyalkanoate monomer synthase operably linked to a promoter functional in a host cell, wherein the nucleic acid molecule comprises a first DNA segment encoding a first module and a second DNA segment encoding a second module, wherein the DNA segments together encode a polyhydroxyalkanoate monomer synthase, and wherein no more than one DNA segment is derived from the *eryA* gene cluster of *Saccharopolyspora erythraea*.
5. The expression cassette of claim 4 wherein the source of at least one DNA segment is bacterial DNA.
6. A method of providing a polyhydroxyalkanoate synthase, comprising:
 - (a) introducing an expression cassette comprising a DNA molecule encoding a polyhydroxyalkanoate synthase operably linked to a promoter functional in a eukaryotic

cell into the eukaryotic cell, wherein the eukaryotic cell is not of plant origin; and

- (b) expressing the DNA molecule encoding the polyhydroxyalkanoate synthase in the eukaryotic cell.

7. The method of claim 6 wherein the polyhydroxyalkanoate synthase is polyhydroxybutyrate synthase.
8. The method of claim 6 wherein the polyhydroxyalkanoate synthase is derived from a bacterium.
9. The method of claim 8 wherein the bacterium is *Alcaligenes eutrophus*.
10. The method of claim 6 wherein the eukaryotic cell is of insect origin.
11. The method of claim 10 wherein the expression cassette is a baculovirus expression cassette.
12. The method of claim 6 further comprising isolating polyhydroxyalkanoate synthase from the eukaryotic cell.
13. A method of providing a polyhydroxyalkanoate polymer, comprising:
 - (a) introducing into a eukaryotic cell (i) a first expression cassette comprising a DNA segment encoding a fatty acid synthase in which the dehydrase activity is inactivated that is operably linked to a promoter functional in the eukaryotic cell, and (ii) a second expression cassette comprising a DNA segment encoding a

polyhydroxyalkanoate synthase operably linked to a promoter functional in the eukaryotic cell; and

- (b) expressing the DNA segments so as to yield a polyhydroxyalkanoate polymer in the eukaryotic cell.

14. The method of claim 13 wherein the eukaryotic cell is of insect origin.
15. The method of claim 13 wherein the dehydrase activity is inactivated by mutating the catalytic site.
16. The method of claim 13 wherein the fatty acid synthase is a rat fatty acid synthase.
17. The method of claim 13 wherein the polyhydroxyalkanoate synthase is a polyhydroxybutyrate synthase.
18. The method of claim 13 wherein the fatty acid synthase produces a premature termination product.
19. The method of claim 13 wherein the fatty acid synthase catalyzes the synthesis of D(-)-3-hydroxybutyrate in the eukaryotic cell.
20. The method of claim 13 wherein the polyhydroxyalkanoate polymer is polyhydroxybutyrate.
21. The method of claim 13 wherein the first and second expression cassettes are on different DNA molecules.

22. An isolated and purified DNA molecule comprising a first DNA segment encoding a first module and a second DNA segment encoding a second module, wherein the DNA segments together encode a recombinant polyhydroxyalkanoate monomer synthase, and wherein no more than one DNA segment is derived from the *eryA* gene cluster of *Saccharopolyspora erythraea*.
23. The isolated DNA molecule of claim 22 wherein the first DNA segment is derived from the *vep* gene cluster of *Streptomyces venezuelae*.
24. The isolated DNA molecule of claim 22 wherein the second DNA segment is derived from the *tyl* gene cluster of *Streptomyces*.
25. The isolated DNA molecule of claim 22 wherein the second DNA segment comprises a DNA encoding a thioesterase which is located at the 3' end of the second DNA segment.
26. The isolated DNA molecule of claim 25 wherein the second DNA segment comprises a DNA encoding an acyl carrier protein which is located 5' to the DNA encoding the thioesterase.
27. The isolated DNA molecule of claim 26 wherein the second DNA segment comprises a DNA encoding a linker region, wherein the DNA encoding the linker region is located between the DNA encoding the acyl carrier protein and the DNA encoding the thioesterase.

28. The isolated DNA molecule of claim 22 wherein the first DNA segment comprises DNA encoding two acyl transferases, wherein the DNA encoding the first acyl transferase is 5' to the DNA encoding the second acyl transferase.
29. The isolated DNA molecule of claim 28 wherein the second acyl transferase adds acyl groups to malonylCoA.
30. The isolated DNA molecule of claim 22 wherein the first DNA segment comprises a DNA encoding a dehydrase.
31. The isolated DNA molecule of claim 22 wherein the first DNA segment comprises a DNA encoding a dehydrase and an enoyl reductase.
32. The isolated DNA molecule of claim 22 wherein the second DNA segment comprises a DNA encoding an inactive dehydrase.
33. The isolated DNA molecule of claim 22 wherein the first DNA segment comprises a DNA encoding an acyl transferase.
34. The isolated DNA molecule of claim 33 wherein the acyl transferase domain binds an acyl CoA substrate.
35. The isolated DNA molecule of claim 22 comprising a first DNA segment encoding a first module and a second DNA segment encoding a second module, wherein the DNA segments together encode a recombinant polyhydroxyalkanoate monomer synthase, and wherein no more than one DNA

segment is derived from the *eryA* gene cluster of *Saccharopolyspora erythraea*.

36. The isolated DNA molecule of claim 35 wherein the first DNA segment encodes the first module from the *vep* gene cluster and the second DNA segment encodes module 7 from the *tyl* gene cluster.
37. A method of providing a polyhydroxyalkanoate monomer, comprising:
 - (a) introducing into a host cell a DNA molecule comprising a DNA segment encoding a recombinant polyhydroxyalkanoate monomer synthase operably linked to a promoter functional in the host cell, wherein the recombinant polyhydroxyalkanoate monomer synthase comprises a first module and a second module; and
 - (b) expressing the DNA encoding the recombinant polyhydroxyalkanoate monomer synthase in the host cell so as to generate a polyhydroxyalkanoate monomer.
38. A method of providing a polyhydroxyalkanoate polymer, comprising:
 - (a) introducing into a host cell a first DNA molecule comprising a DNA segment encoding a recombinant polyhydroxyalkanoate monomer synthase operably linked to a promoter functional in a host cell, wherein the recombinant polyhydroxyalkanoate monomer synthase comprises a first module and a second module;
 - (b) introducing into the host cell of step (a) a second DNA molecule comprising a DNA segment encoding a

polyhydroxyalkanoate synthase operably linked to a promoter functional in the host cell; and

- (c) expressing the DNAs encoding the recombinant polyhydroxyalkanoate monomer synthase and polyhydroxyalkanoate synthase in the host cell so as to generate a polyhydroxyalkanoate polymer.

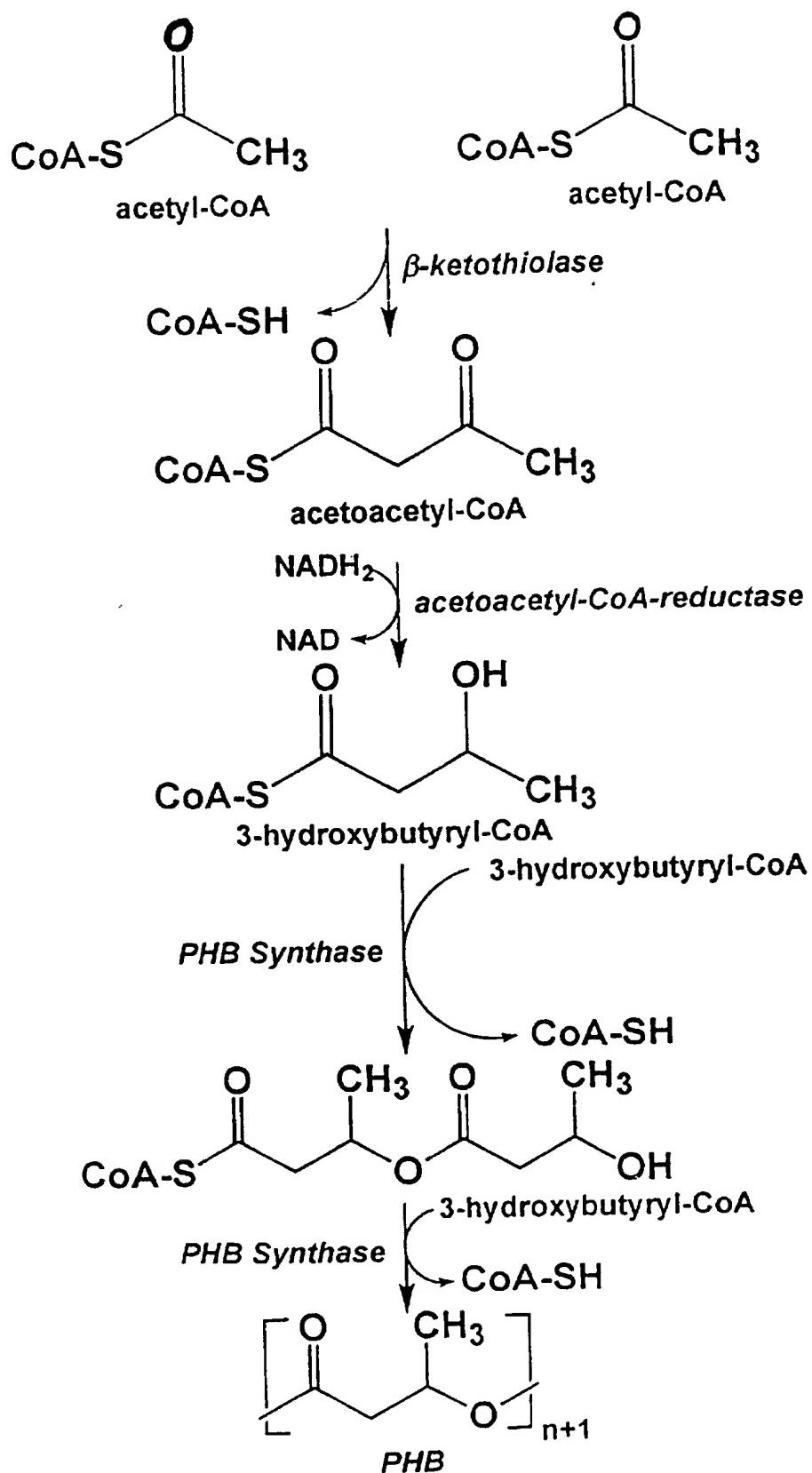
39. The method of claim 37 or 38 wherein the first DNA segment encodes the first module from the *vep* gene cluster and the second DNA segment encodes module 7 from the *tyl P* gene cluster.
40. An isolated and purified DNA molecule comprising a first DNA segment encoding a fatty acid synthase and a second DNA segment encoding a module of a polyketide synthase.
41. The isolated DNA molecule of claim 40 wherein the second DNA segment encodes a β -ketoacyl synthase amino-terminal to an acyltransferase which is amino-terminal to a ketoreductase which is amino-terminal to an acyl carrier protein which is amino-terminal to a thioesterase.
42. The isolated DNA molecule of claim 40 wherein the second DNA segment is 3' to the DNA encoding the fatty acid synthase.
43. The isolated DNA molecule of claim 40 wherein the second DNA segment is separated from the first DNA segment by a DNA encoding a linker region.

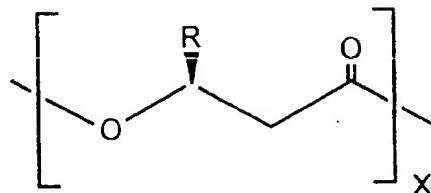
44. The isolated DNA molecule of claim 41 wherein the DNA encoding the linker region is selected from the group consisting of *tyl* ORF1 ACP₁-KS₂, *tyl* ORF1 ACP₂-KS₃, *tyl* ORF3 ACP₅-KS₆, *eryA* ORF1 ACP₁-KS₁, *eryA* ORF1 ACP₂-KS₂, *eryA* ORF2 ACP₃-KS₄, and *eryA* ORF2 ACP₅-KS₆.
45. The isolated DNA molecule of claim 40 comprising a first DNA segment encoding a fatty acid synthase and a second DNA segment encoding a module of a polyketide synthase.
46. A method of providing a polyhydroxyalkanoate monomer, comprising:
 - (a) introducing into a host cell a DNA molecule comprising a first DNA segment encoding a fatty acid synthase and a second DNA segment encoding a polyketide synthase, wherein the first DNA segment is 5' to the second DNA segment, wherein the first DNA segment is operably linked to a promoter functional in the host cell, and wherein the first DNA segment is linked to the second DNA segment so that the linked DNA segments express a fusion protein; and
 - (b) expressing the DNA molecule in the host cell so as to generate a polyhydroxyalkanoate monomer.
47. The method of claim 46 wherein the host cell is selected from the group consisting of insect cells, *Streptomyces* cells and *Pseudomonas* cells.
48. The method of claim 46 wherein the DNA encoding the fatty acid synthase is eukaryotic in origin.

49. The method of claim 46 wherein the DNA encoding the fatty acid synthase is prokaryotic in origin.
50. The method of claim 46 wherein the DNA encoding the polyketide synthase module is derived from DNA encoding the *tyl* module F.
51. An expression cassette comprising a DNA molecule comprising a DNA segment encoding a fatty acid synthase and a polyhydroxyalkanoate synthase.
52. A method of providing a polyhydroxyalkanoate monomer synthase, comprising:
 - (a) introducing an expression cassette comprising a DNA molecule encoding a polyhydroxyalkanoate synthase operably linked to a promoter functional in a host cell, wherein the DNA comprises a first DNA segment encoding a first module and a second DNA segment encoding a second module wherein the DNA segments together encode a polyhydroxyalkanoate monomer synthase; and
 - (b) expressing the DNA molecule in the host cell.
53. An isolated and purified DNA molecule comprising a DNA segment encoding a *Streptomyces venezuelae* polyketide synthase.
54. The isolated DNA molecule of claim 53 wherein the DNA segment comprises SEQ ID NO:1.

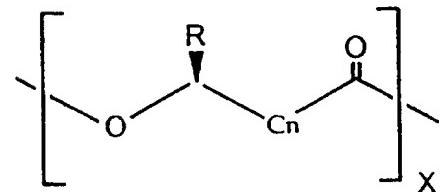
55. The isolated DNA molecule of claim 53 wherein the DNA segment encodes a polypeptide having an amino acid sequence comprising SEQ ID NO:2.
56. The expression cassette of claim 4 wherein the first DNA segment encodes the first module from the *vep* gene cluster and the second DNA segment encodes module 7 from the *tyl P* gene cluster.
57. The expression cassette of claim 4 further comprising a third DNA segment encoding a polyhydroxyalkanoate synthase.
58. The method of claim 37 wherein the DNA molecule further comprises a DNA segment encoding a polyhydroxyalkanoate synthase.
59. The isolated DNA molecule of claim 22 or 36 further comprising a DNA segment encoding a polyhydroxyalkanoate synthase.
60. The method of claim 53 wherein the expression cassette further comprises a second DNA molecule encoding a polyhydroxyalkanoate synthase

Figure 1





<u>R-group</u>	<u>Monomer</u>	<u>Abbreviation</u>
methyl	3-hydroxybutyrate	(3HB)
ethyl	3-hydroxyvalerate	(3HV)
propyl	3-hydroxycaproate	(3HC)
butyl	3-hydroxyheptanoate	(3HH)
pentyl	3-hydroxyoctanoate	(3HO)
hexyl	3-hydroxynonanoate	(3HN)
heptyl	3-hydroxydecanoate	(3HD)
octyl	3-hydroxyundecanoate	(3HUD)
nonyl	3-hydroxydodecanoate	(3HDD)



- n = 1 3-hydroxyacyl monomer
 n = 2 4-hydroxyacyl monomer
 n = 3 5-hydroxyacyl monomer

Figure 2

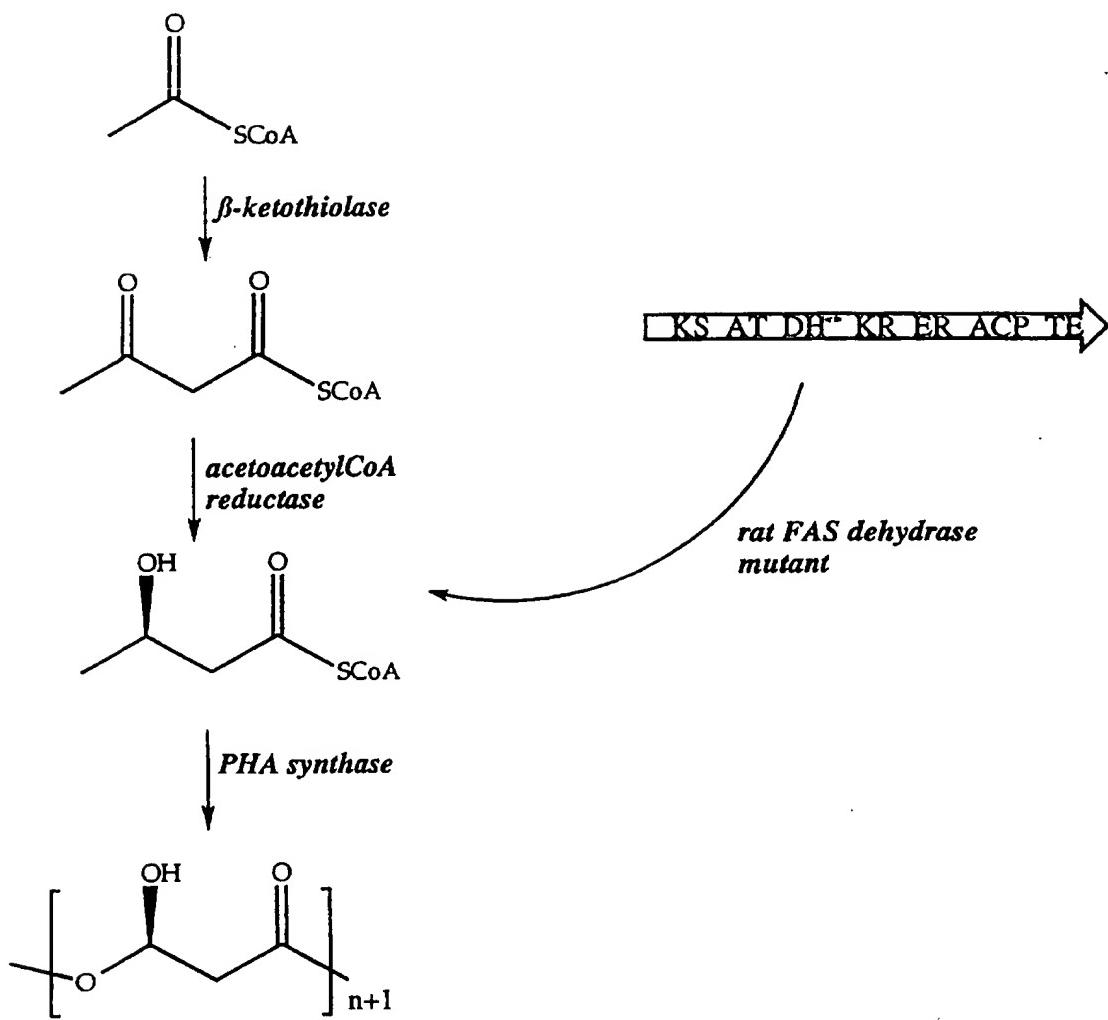


Figure 3

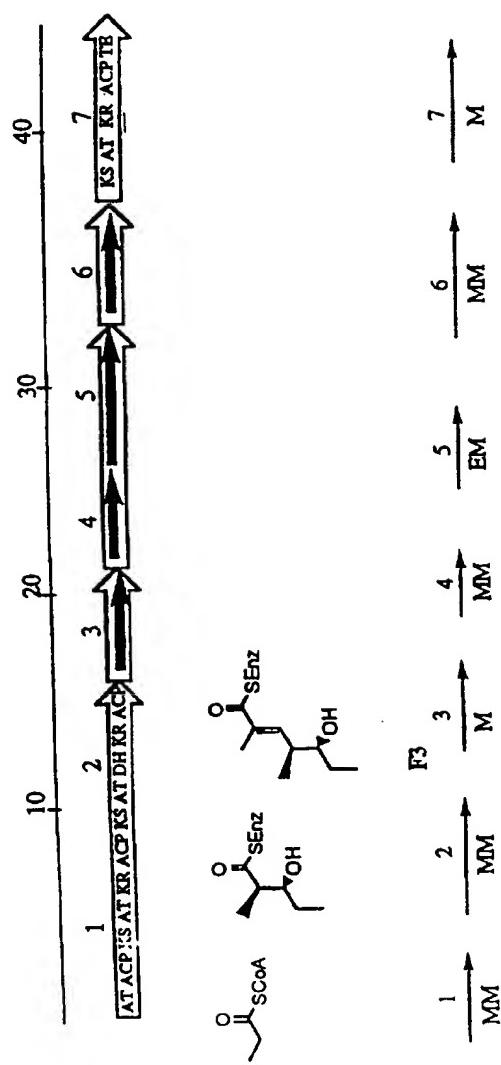


Figure 4

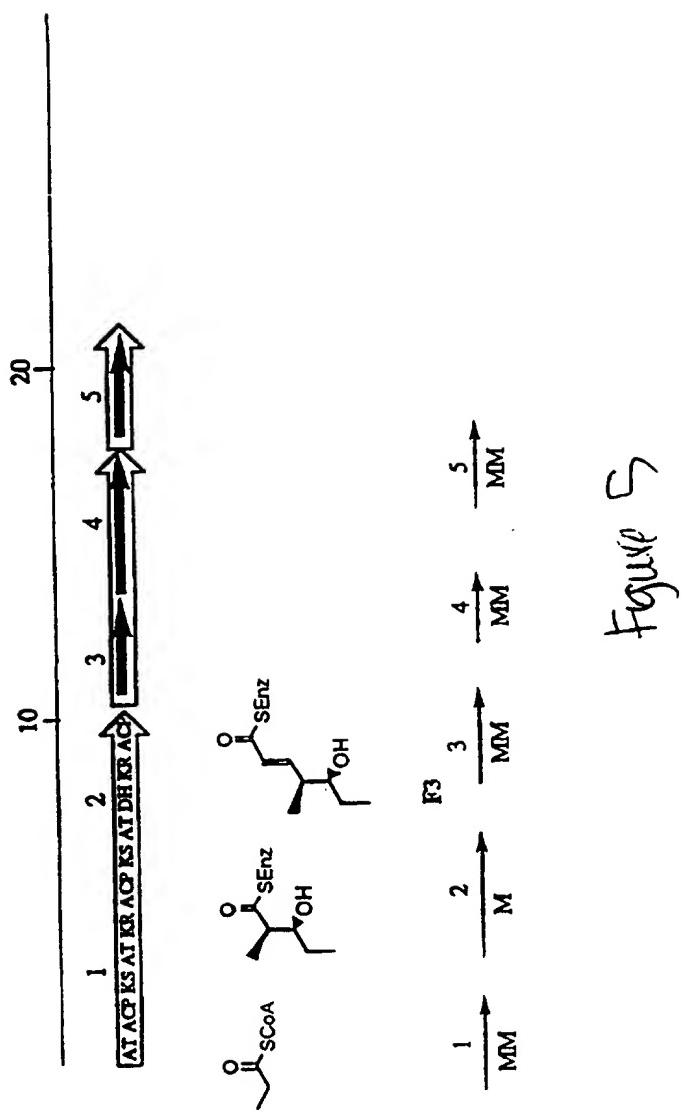


Figure 5

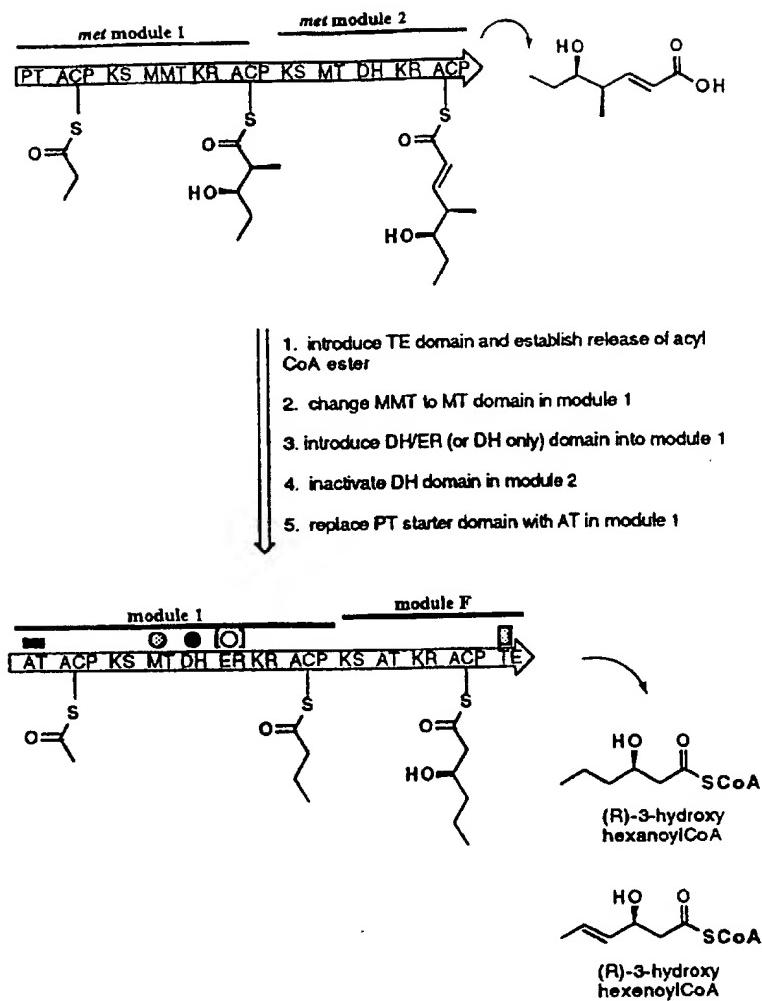


Fig. 6

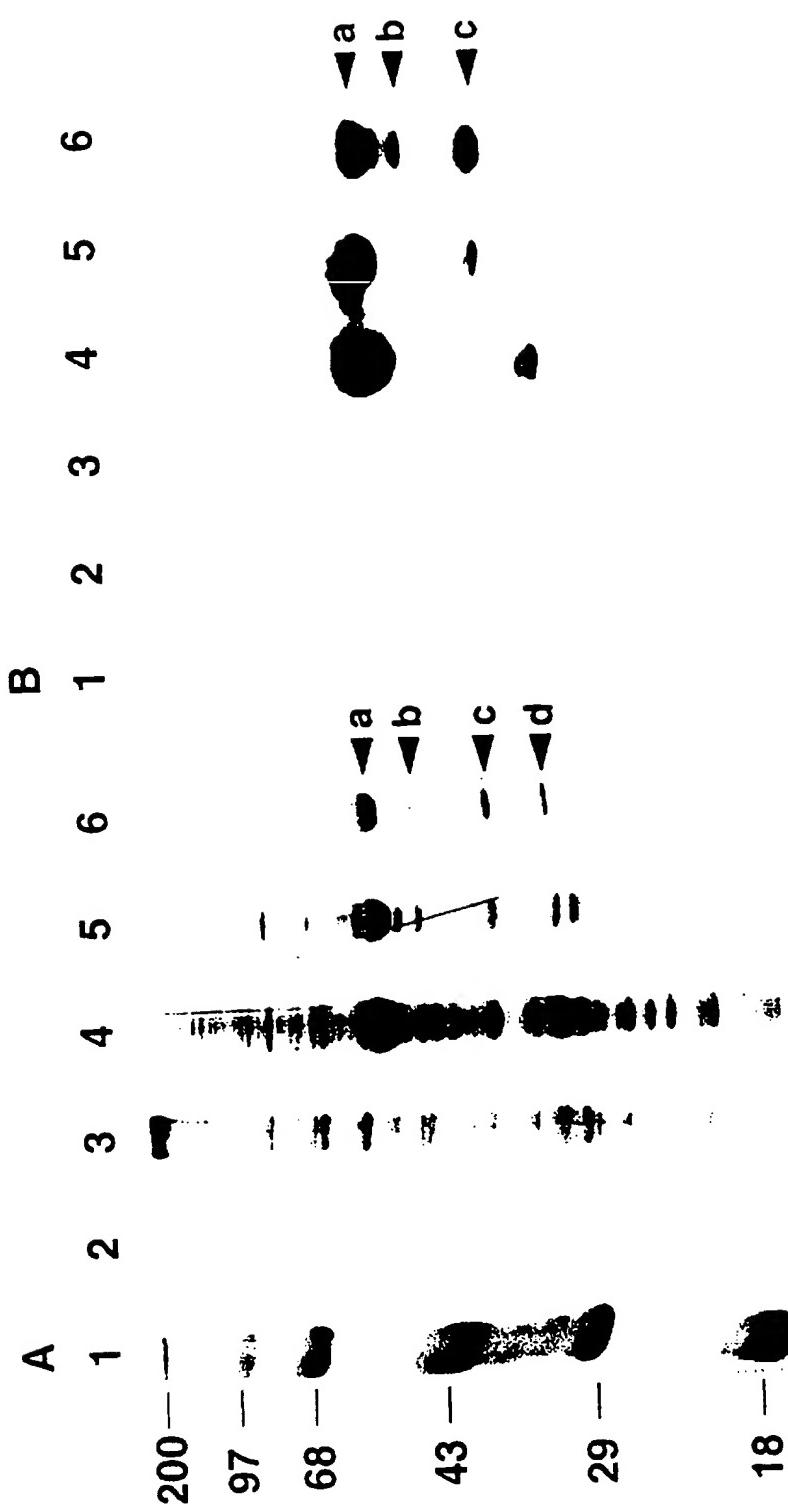


Figure 7

Figure 8

N-terminal sequence determined for PHA synthase

	1	10	20	25
a	MATGKGAA	ASTQEGKSQ	PFKVTPGP	-
b	AAASTQEGKSQPFKVTPGP-			
c	STQEGKSQPFKVTPGP-			

Figure 9

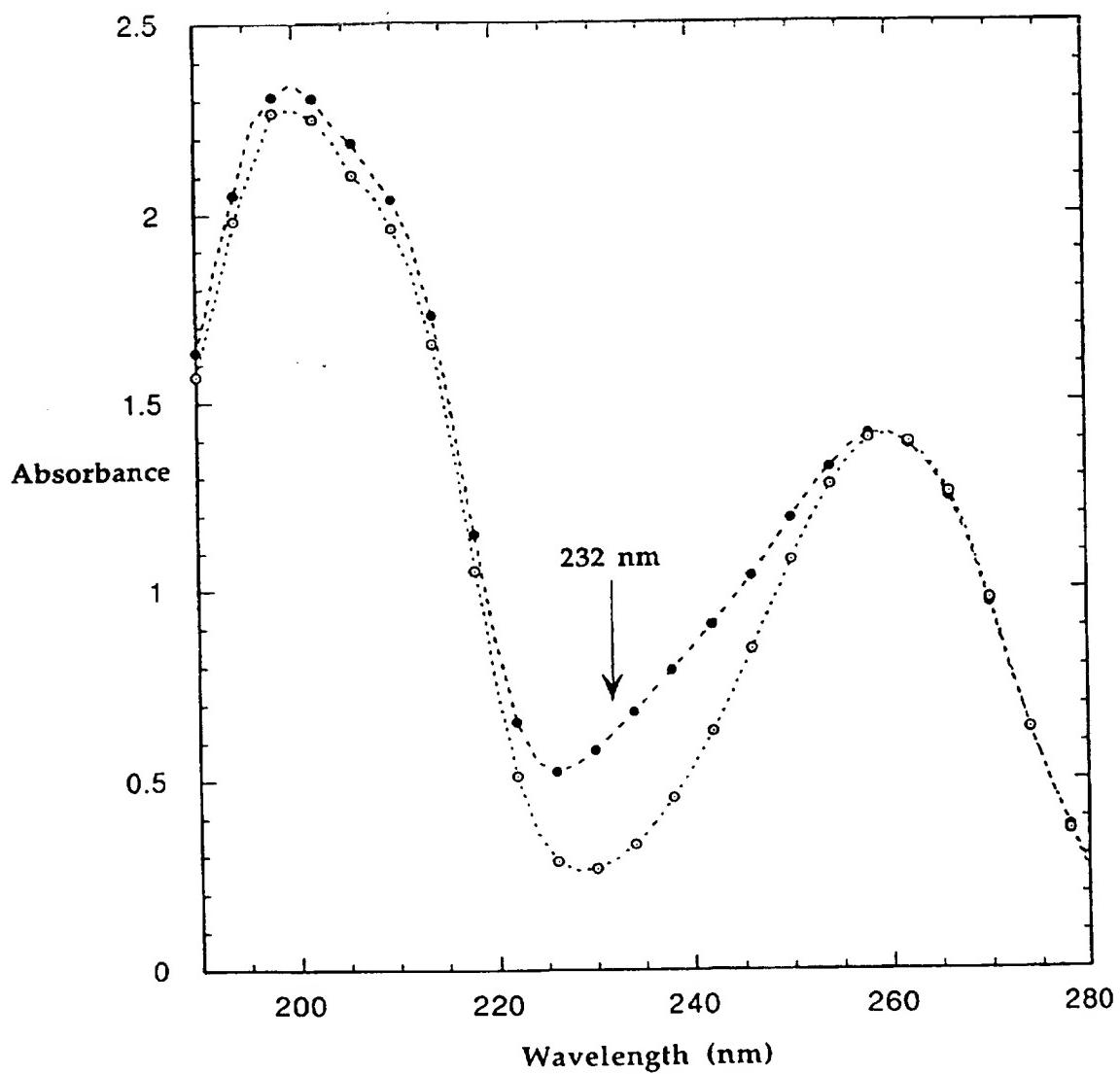


Figure 10

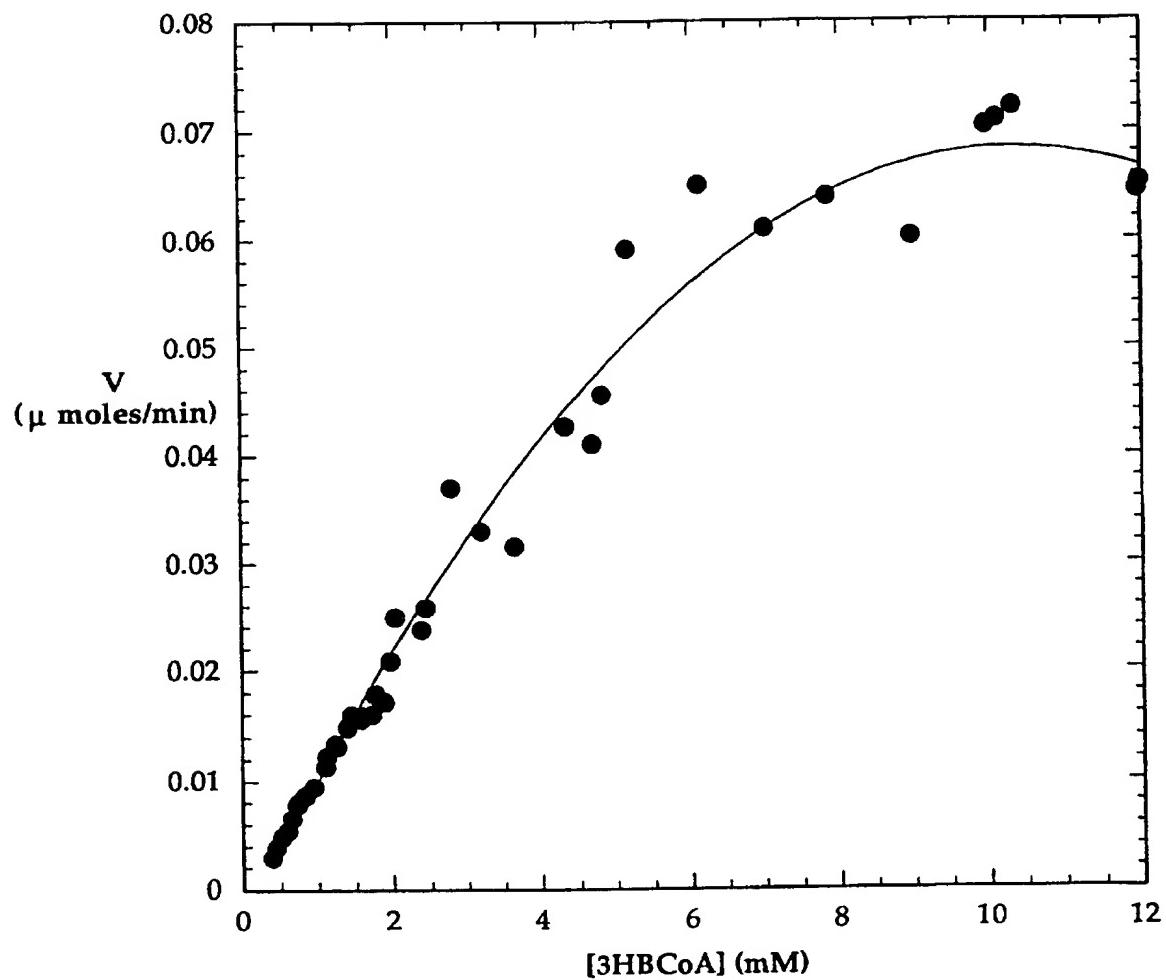


Figure 11

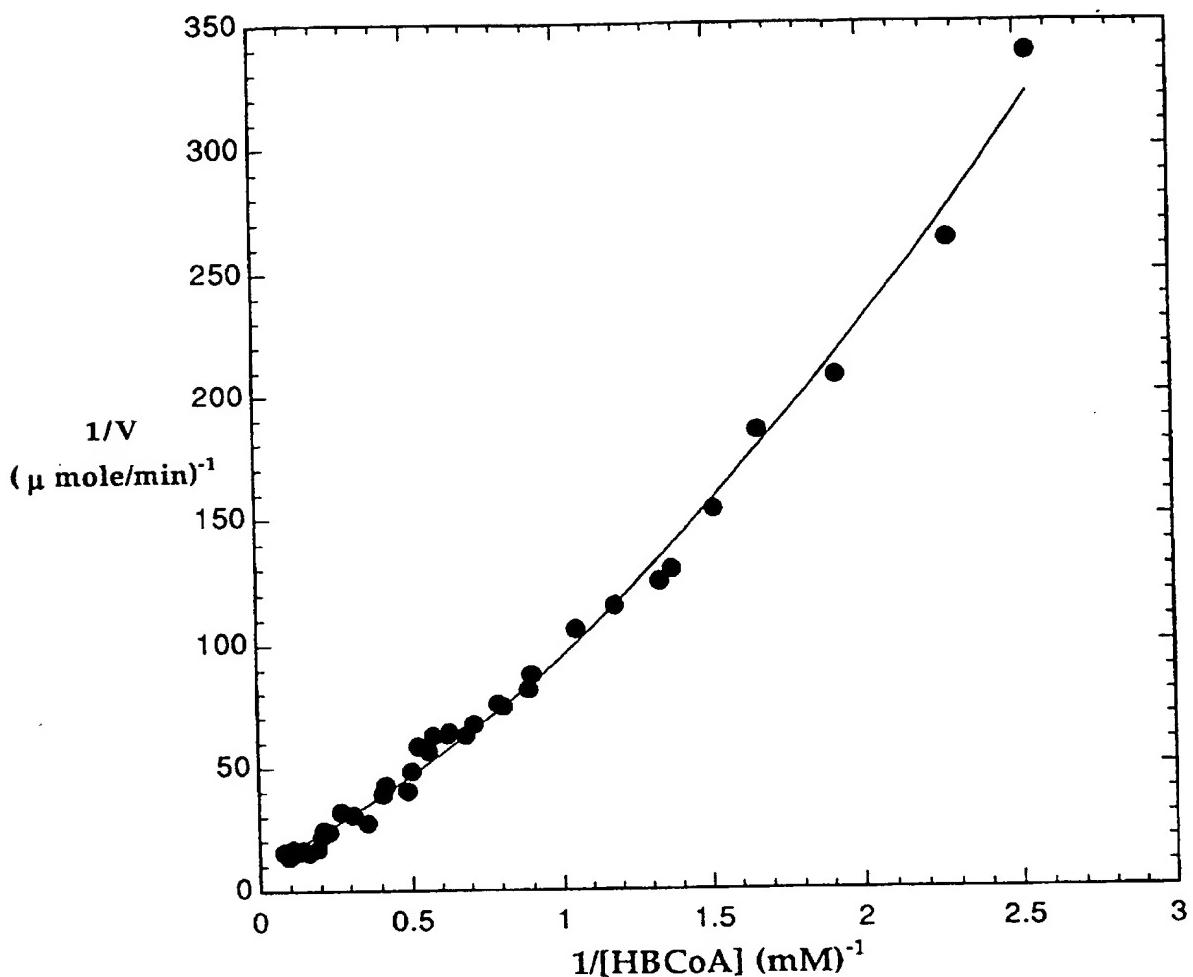


Figure 18

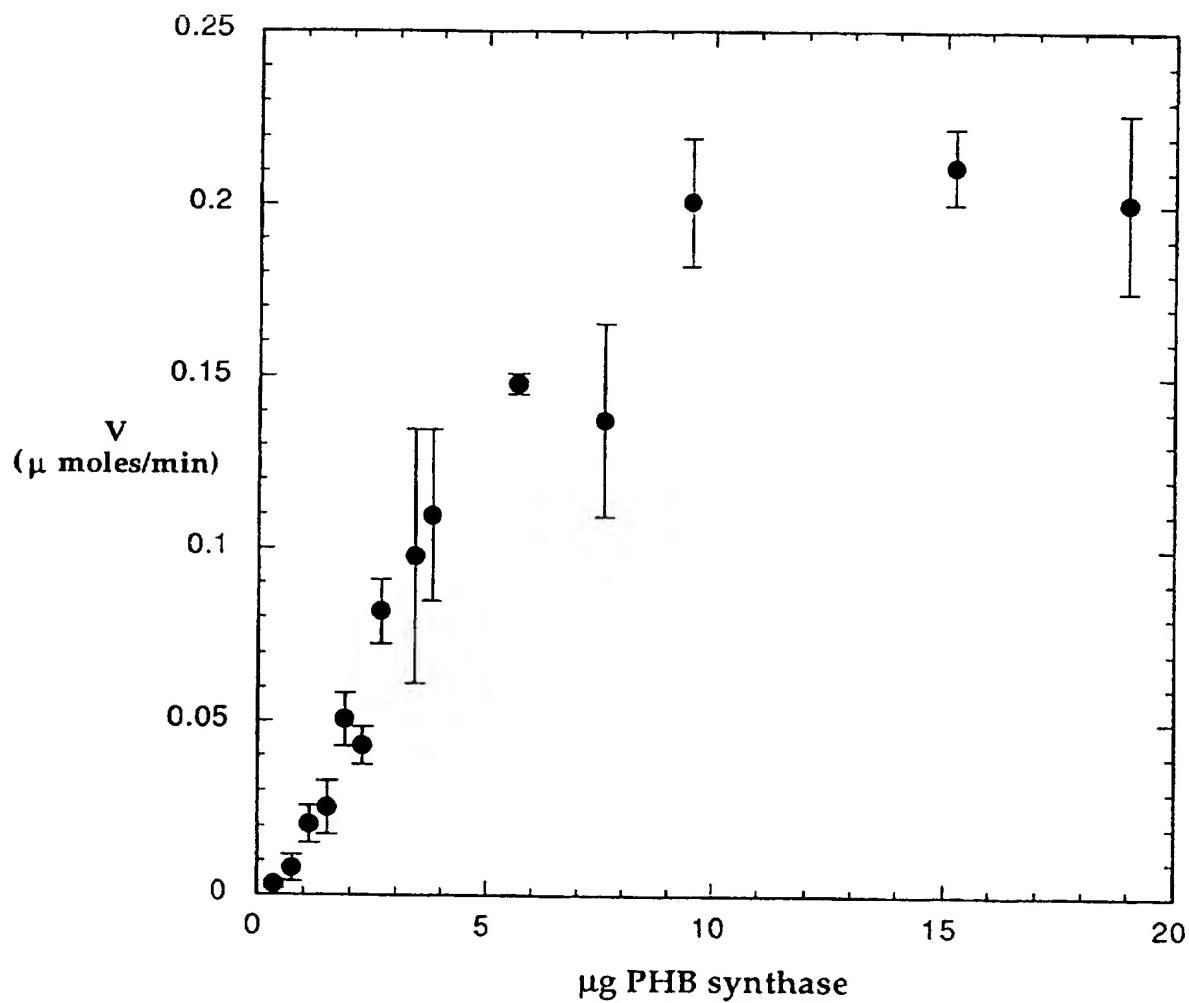


Figure 13

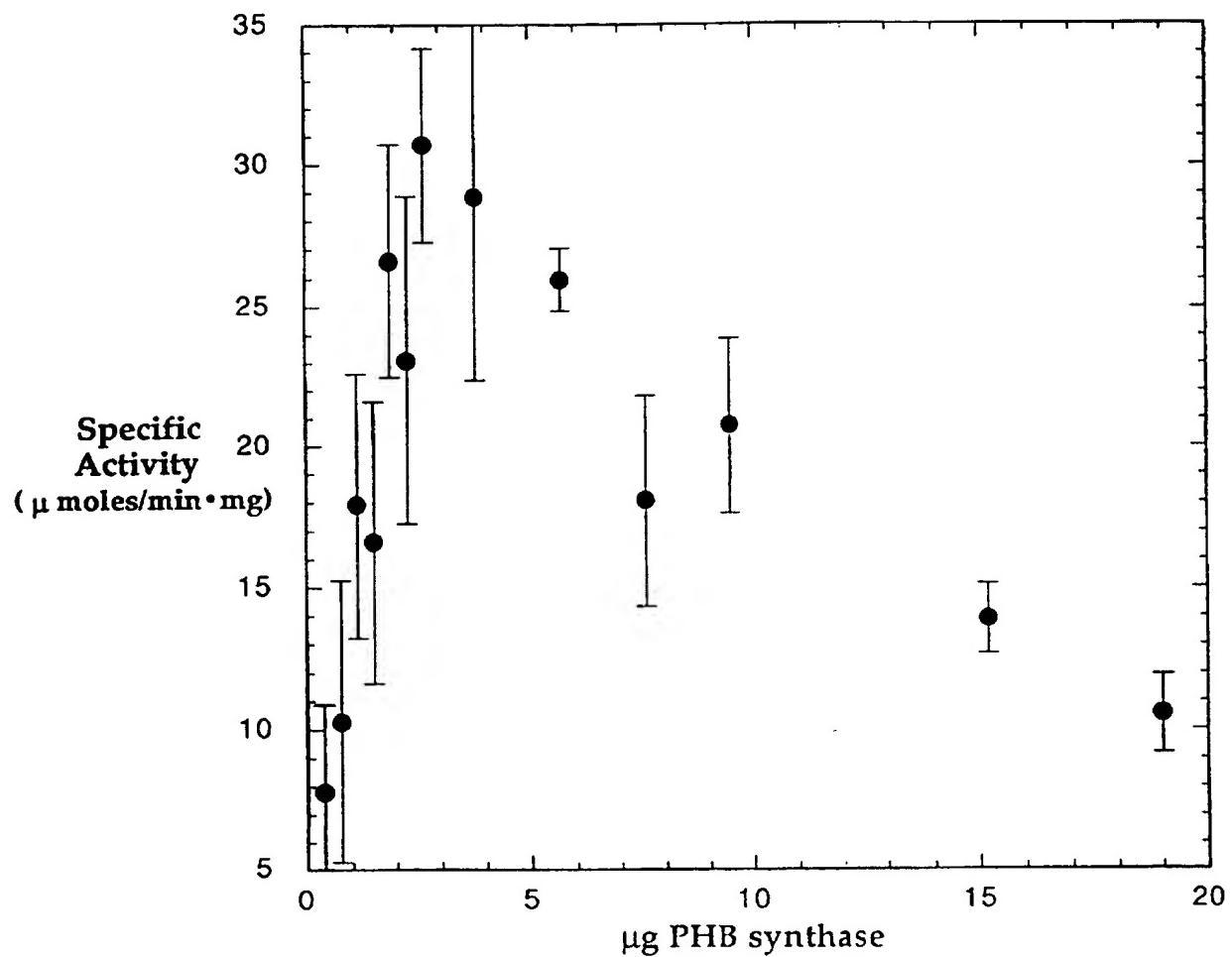


Figure 14

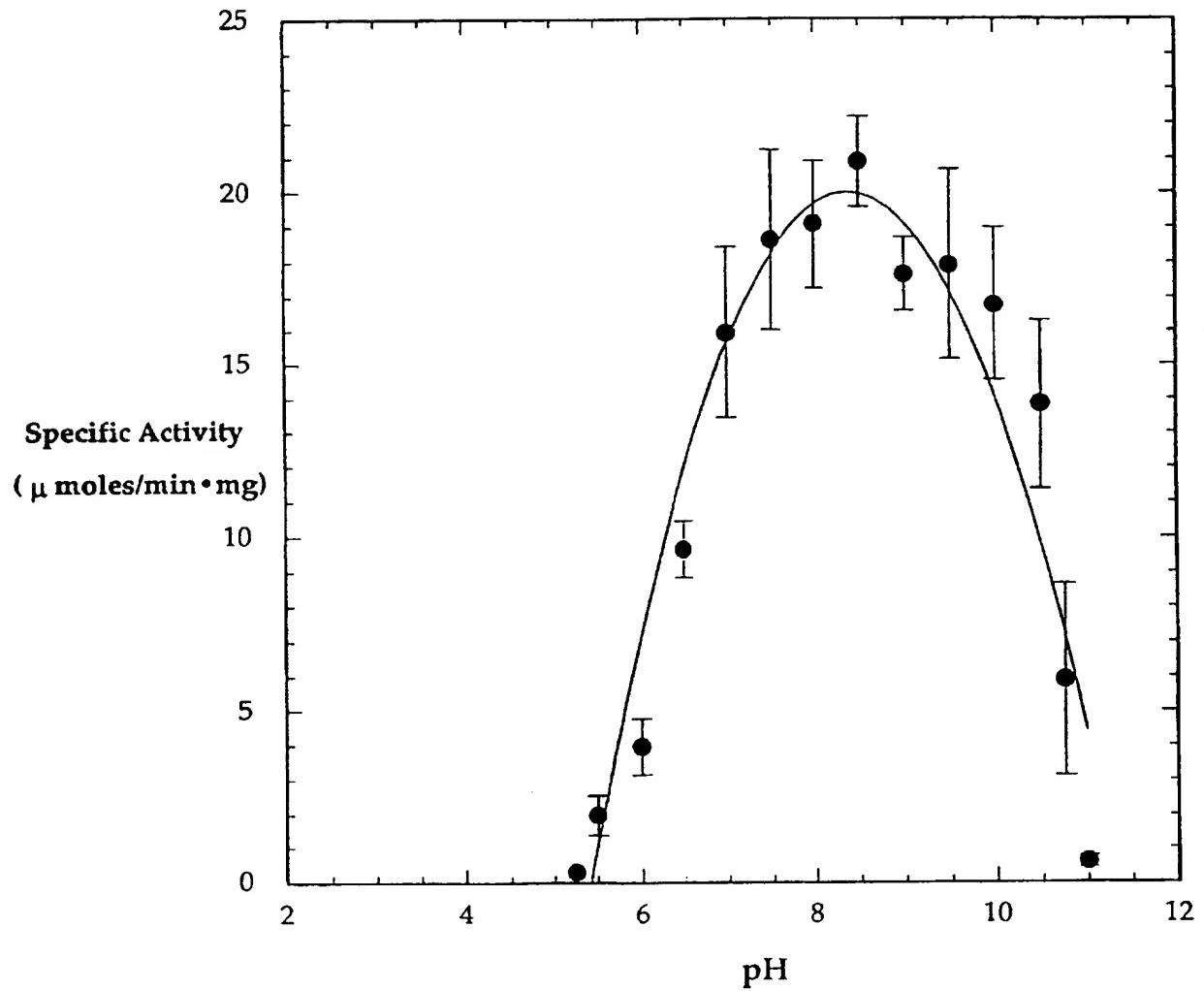
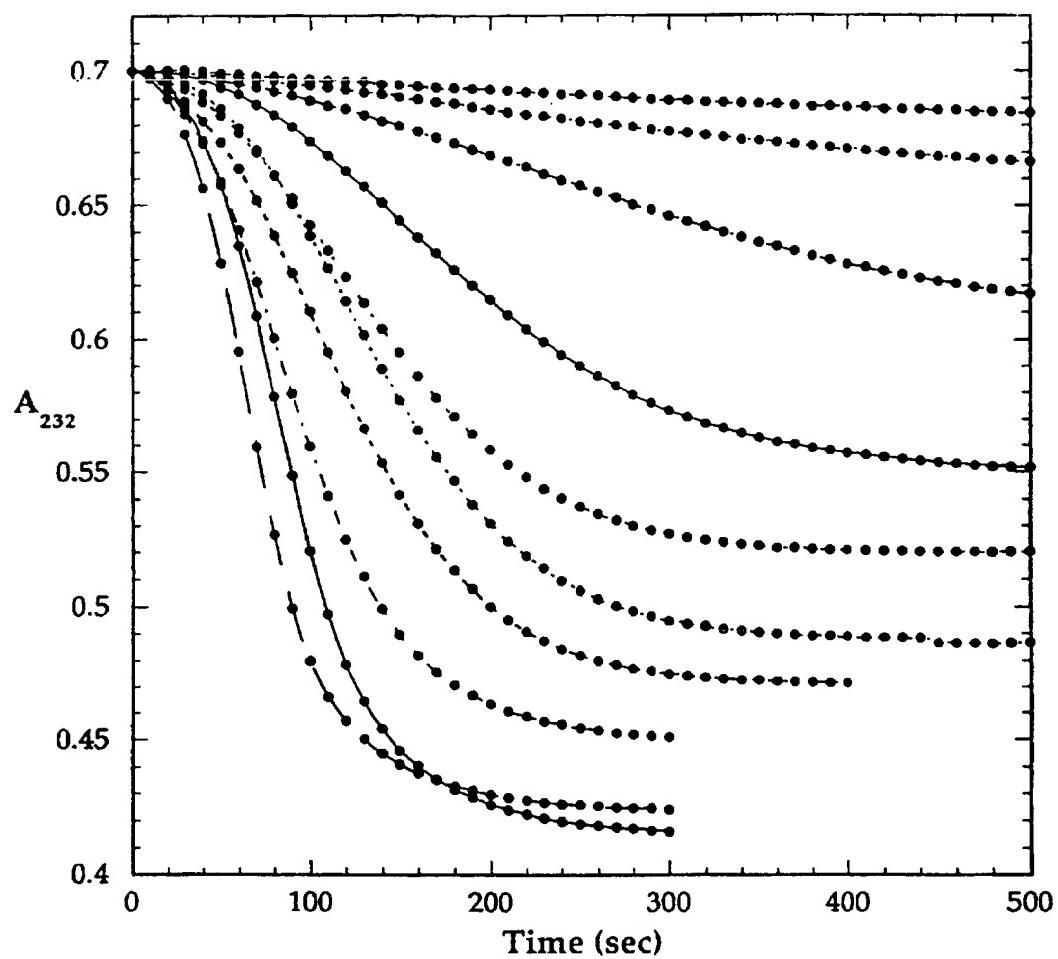


Figure 15



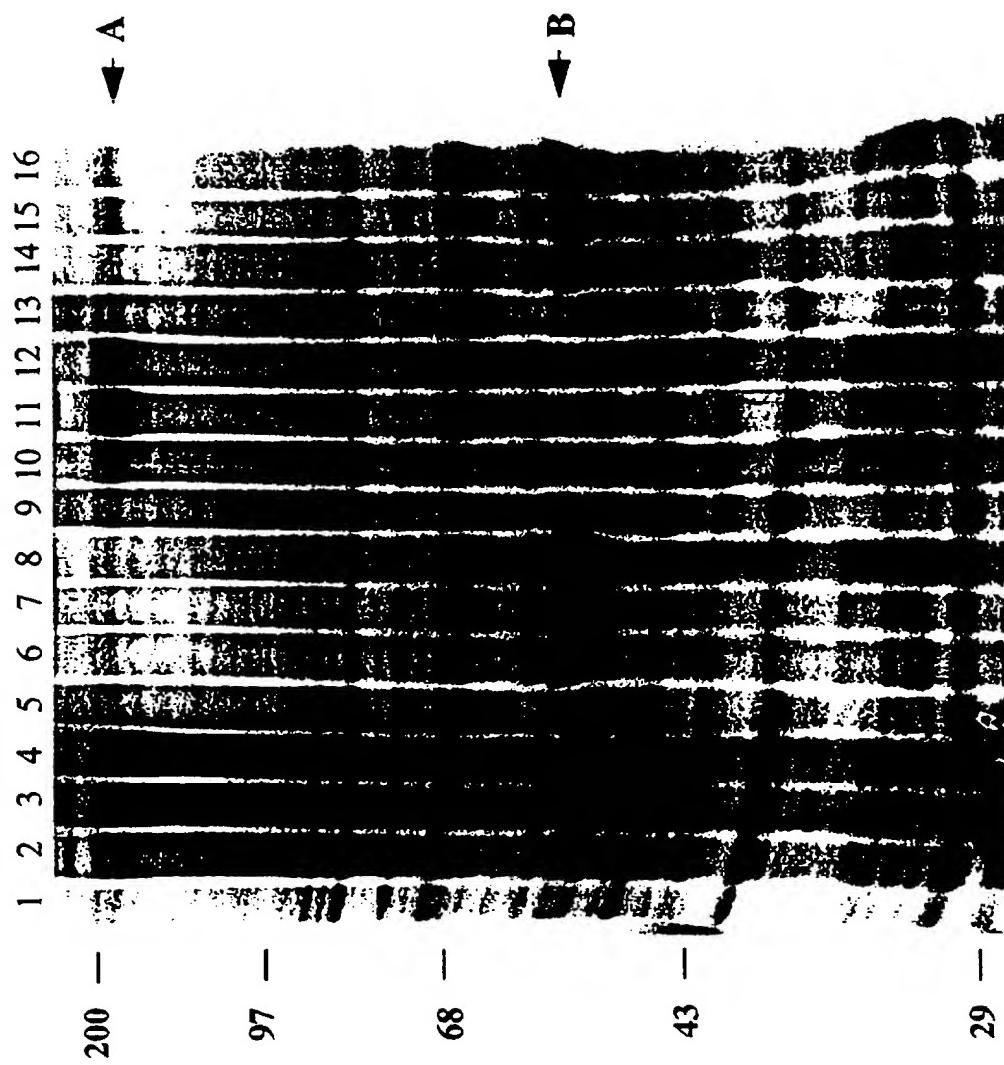


Figure 1b

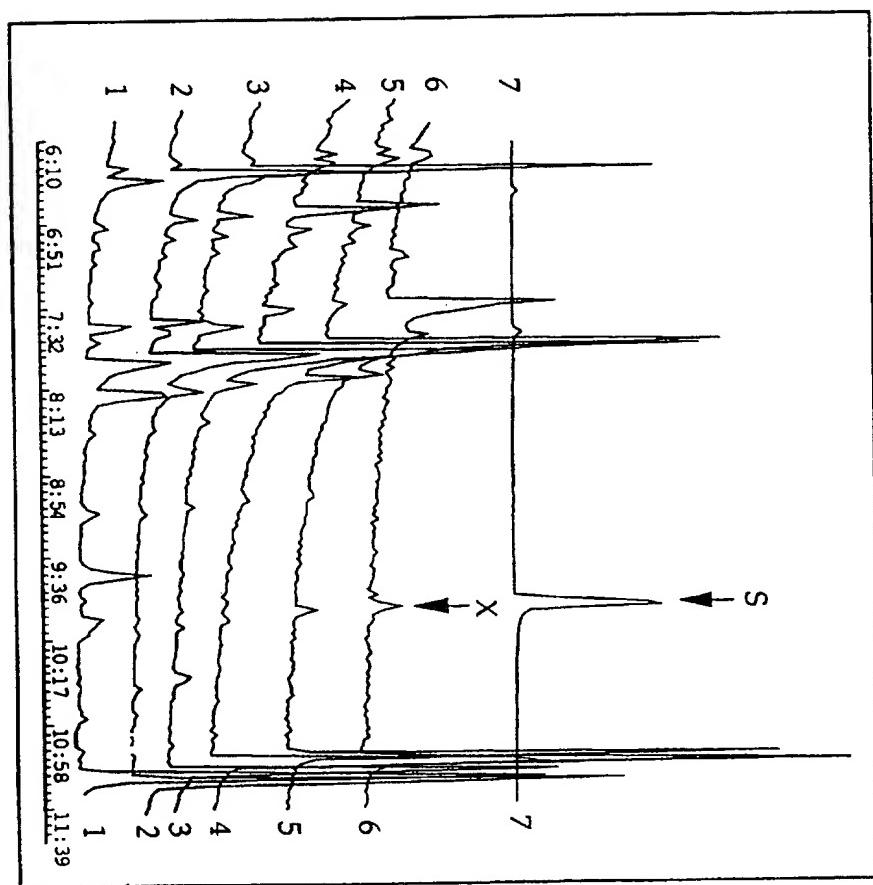


Figure 7

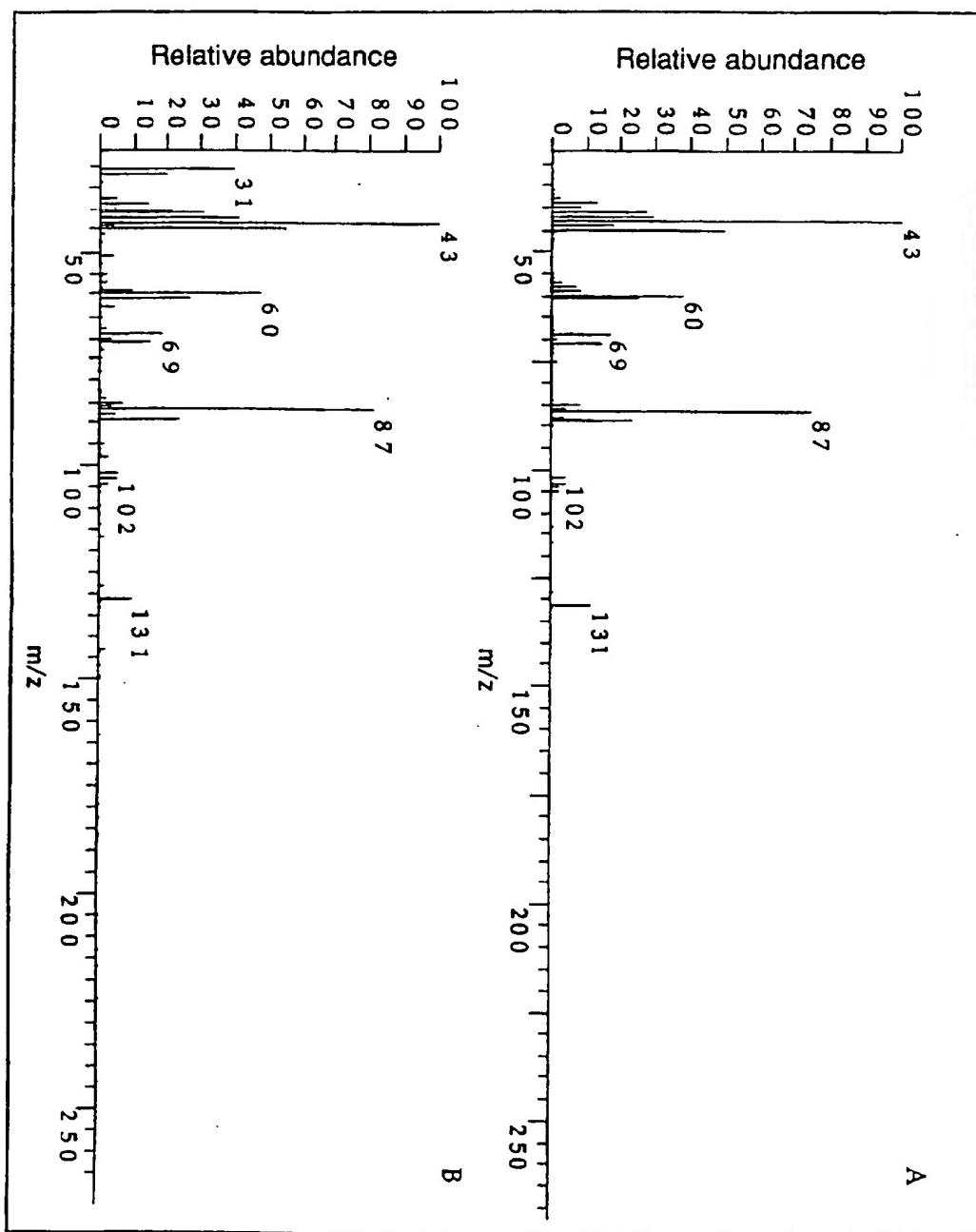


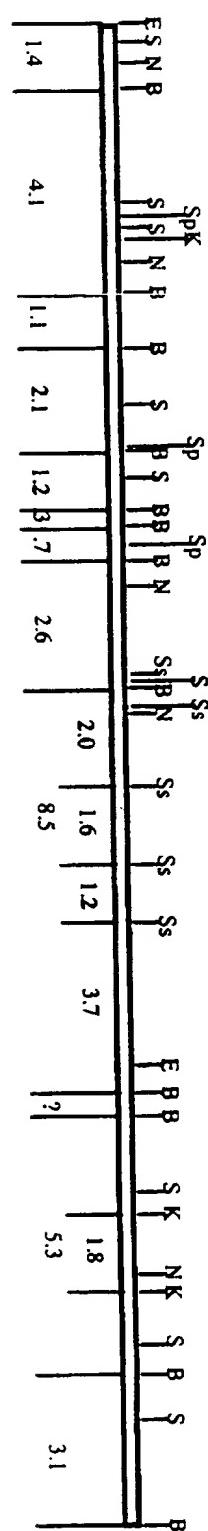
Figure 18

The Map of *vep* Cluster

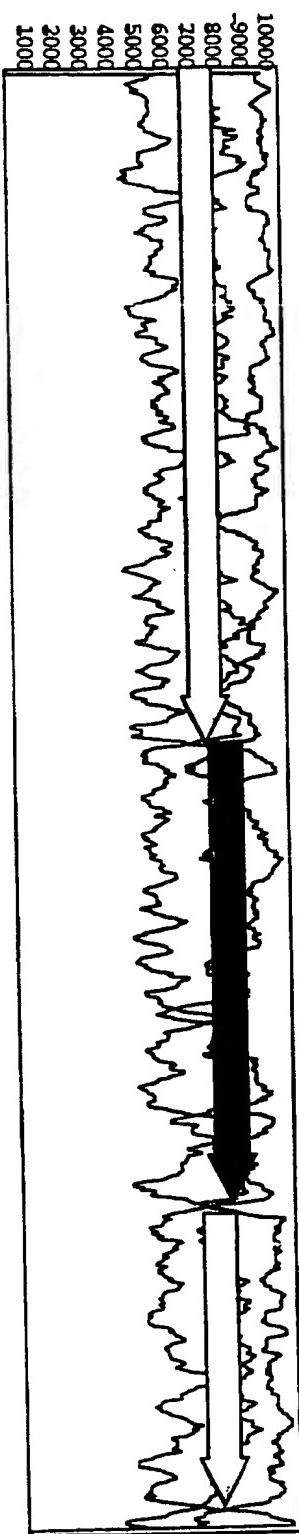
PCT/US96/20119

WO 97/22711

19/33



Open Reading Frame analysis



Modular structure

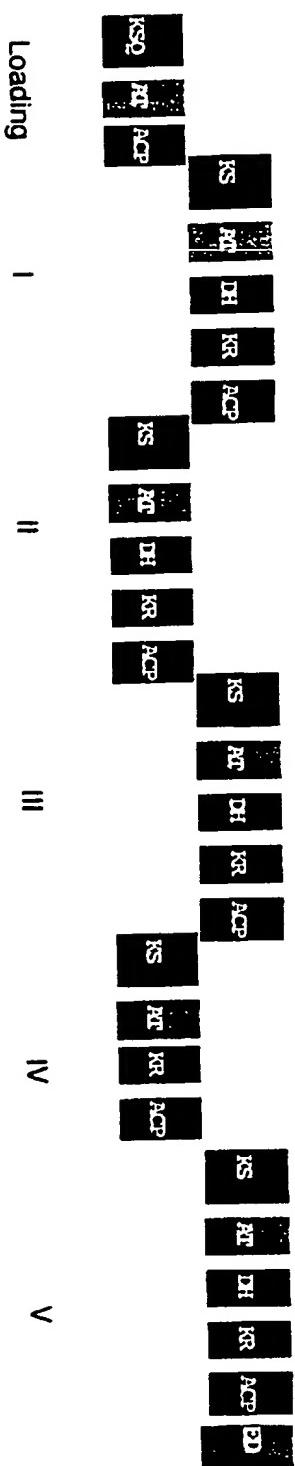


Figure 19

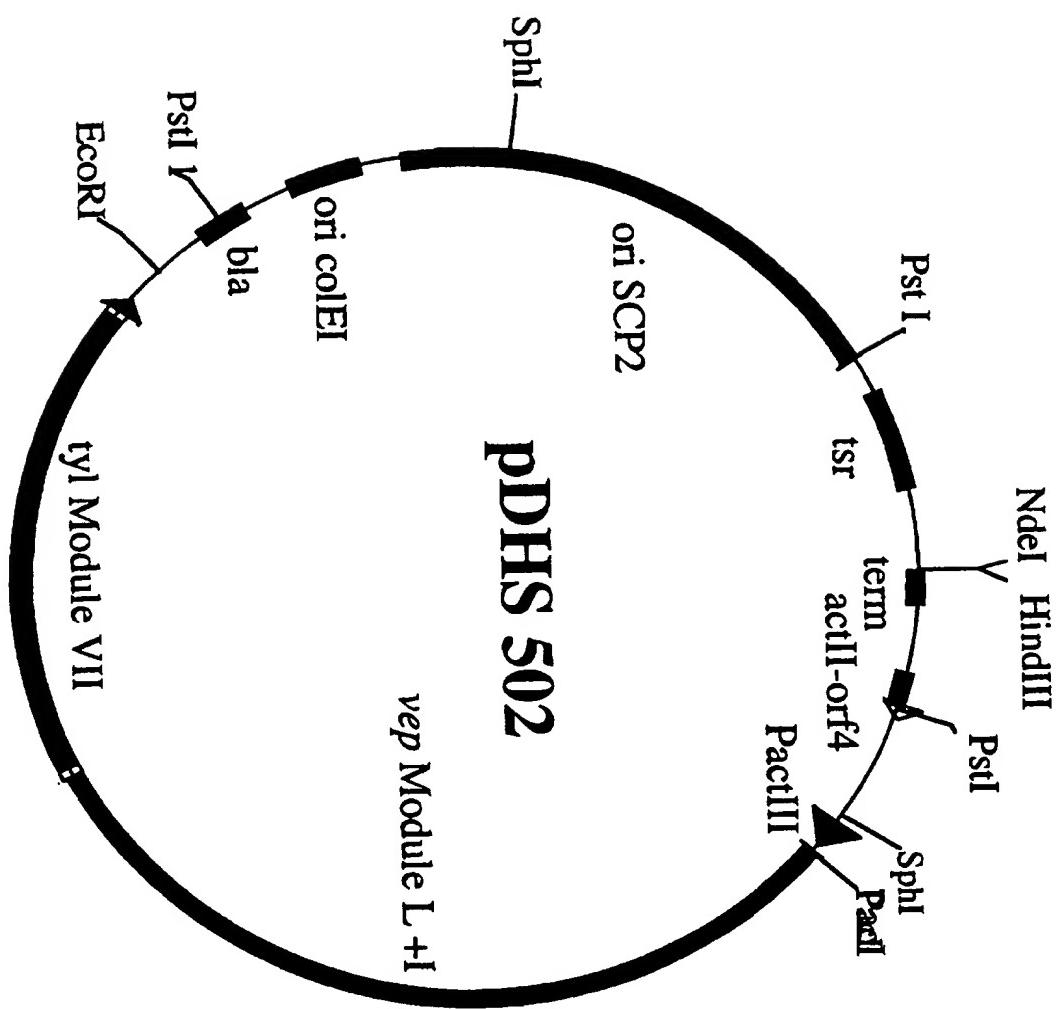


Figure 20

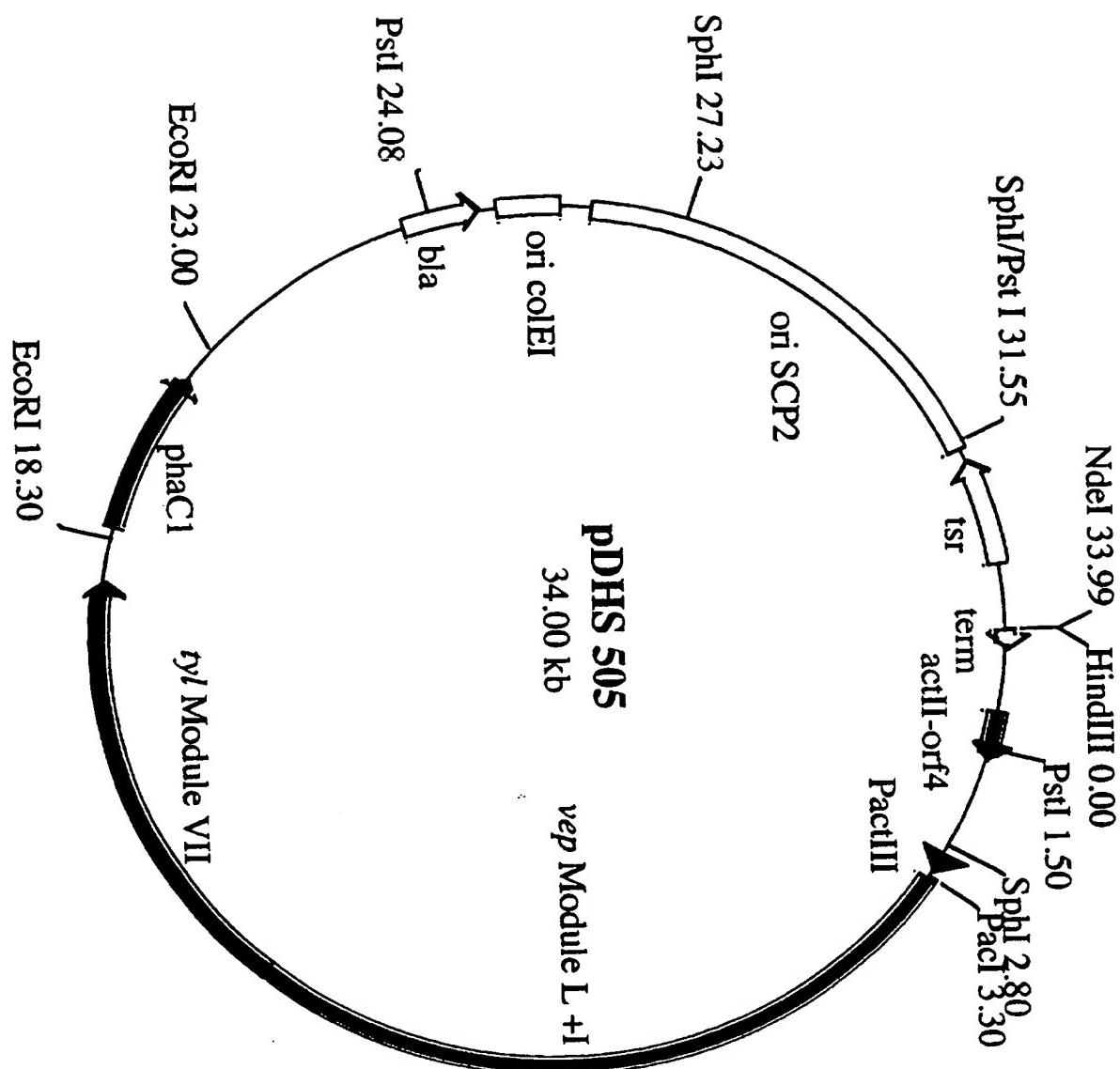
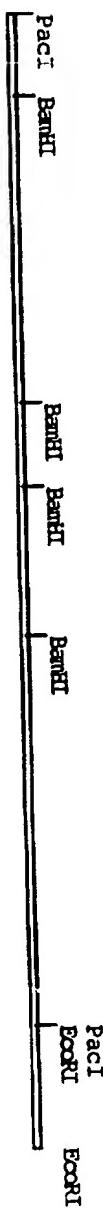


Figure 21

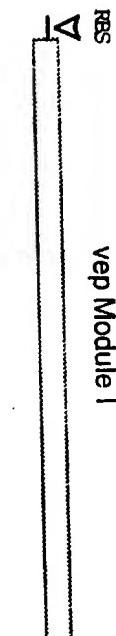
pDHS505 Construction Procedure

Restriction map of pDHS505 insert



PCT/US96/20119

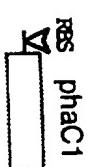
Major steps in the construction



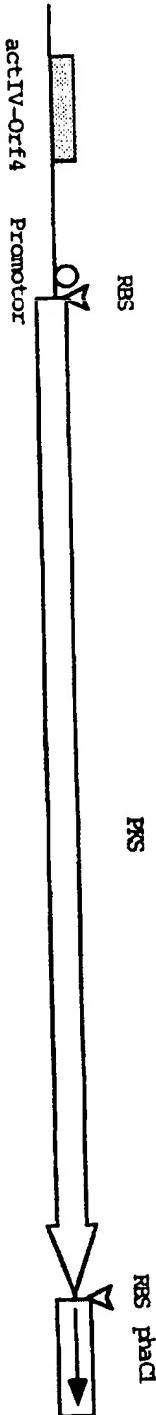
1. A Streptomyces Ribosome Binding Site (RBS) was introduced 6 nucleotides upstream of the translation start site to enhance gene translation in the host.



2. The *tyl* Module VII was recombined by a *Bam*H site with the *vep* Module I to give out a complete polyketide synthase Open Reading Frame (ORF) with a Thioesterase at 3'-end.



3. The *phaC1* gene was transcriptionally coupled with the PKS gene. The second Ribosome Binding Site (RBS) was introduced to facilitate the gene translation.



4. The whole expression construct was put under the control of act promoter and the actIV-Orf4 provides an activator which enhances the transcription and expression of the genes.

Figure 27

23/33

1	TTAATTAAGGAGGACCATC	ATG AAC GAG GCC ATC GCC GTC GTC GGC ATG TCC TGC CGC CTG CCG	64
	M N E A I A V V G M S C R L P	15	
1	AAG GCC TCG AAC CCG GCC GCC TTC TGG GAG CTG CTG CGG AAC GGG GAG AGC GCC GTC ACC	124	
16	K A S N P A A F W E L L R N G E S A V T	35	
65	GAC GTG CCC TCC CGG CGG TGG ACG TCG GTG CTC GGG GGA GCG GAC GCC GAG GAG CCG GCG	184	
36	D V P S G R W T S V L G G A D A E E P A	55	
125	GAG TCC GGT GTC CGC CGG GGC TTC CTC GAC TCC CTC GAC CTC TTC GAC GCG GCC TTC	244	
56	E S G V R R G G F L D S L D L F D A A F	75	
185	TTC GGA ATC TCG CCC CGT GAG GCC GCC ATG GAC CCG CAG CAG CGA CTG GTC CTC GAA	304	
76	F G I S P R E A A A M D P Q Q R L V L E	95	
245	CTC GCC TGG GAG GCG CTG GAG GAC GCC GGA ATC GTC CCC GGC ACC CTC GCC GGA AGC CGC	364	
96	L A W E A L E D A G I V P G T L A G S R	115	
305	ACC GCC GTC TTC GTC GGC ACC CTG CGG GAC GAC TAC ACG AGC CTC CTC TAC CAG CAC GGC	424	
116	T A V F V G T L R D D Y T S L L Y Q H G	135	
365	GAG CAG GCC ATC ACC CAG CAC ACC ATG GCG GGC GTG AAC CGG GGC GTC ATC GCC AAC CGC	484	
136	E Q A I T Q H T M A G V N R G V I A N R	155	
425	GTC TCG TAC CAC CTC GGC CTG CAG GGC CCG AGC CTC ACC GTC GAC GCC GCG CAG TCG TCC	544	
156	V S Y H L G L Q G P S L T V D A A Q S S	175	
485	TG G CTC GTC GGC GTG CAC CTG GCC TGC GAG TCC CTG CGC GCC GGG GAG TCC ACG ACG GCG	604	
176	S L V A V H L A C E S L R A G E S T T A	195	
545	CTC GTC GGC GGC GTG AAC CTC AAC ATC CTC GCG GAG AGC GCC GTG ACC GAG GAG CGC TTC	664	
196	L V A G V N L N I L A E S A V T E E R F	215	
605	GGT GGA CTC TCC CGG GAC GGC ACC GCC TAC ACC TTC GAC GCG CGG GCC AAC GGA TTC GTC	724	
216	G G L S P D G T A Y T F D A R A N G F V	235	
665	CGG GGC GAG GGC GGC GGA GTC GTC GTC CTC AAG CGG CTC TCC CGC GCC CTC GCC GAC GGC	784	
236	R G E G G G V V V L K P L S R A L A D G	255	
725	GAC CGT GTC CAC GGC GTC ATC CGC GCC AGC GCC GTC AAC AAC GAC GGA GCC ACC CCG GGT	844	
256	D R V H G V I R A S A V N N D G A T P G	275	
785	CAC ACC GTG CCC AGC AGG GCC CAG GAG AAG GTG CTG CGC GAG GCG TAC CCG AAG GCG	904	
276	L T V P S R A A Q E K V L R E A Y R K A	295	
845	GCC CTG GAC CCG TCC GCC CAG TAC GTC GAA CTC CAC GGC ACC GGA ACC CCC GTC GGC	964	
296	A L D P S A V Q Y V E L H G T G T P V G	315	
905	GAC CCC ATC GAG GCC GCG CTC GGC GCC TCG GCG CCC CCC GCG GAC GAA	1024	
316	D P I E A A A L G A V L G S A R P A D E	335	
965	CTG CTC GTC GGC TCG GCC AAG ACG AAC GTC GGG CAC CTC GAA GGC GCC GCC ATC	1084	
336	P L L V G S A K T N V G H L E G A A G I	355	
1025	GGC CTC ATC AAG ACG CTC CTC GCG CTC GGC CGG CGC CGG ATC CCG CGG AGC CTC AAC	1144	
356	V G L I K T L L A L G R R R I P A S L N	375	
1085	TTC CGT ACG CCC CAC CGG GAC ATC CCG CTC GAC ACC CTC GGG CTC GAC GTG CCC GAC GGC	1204	
376	F R T P H P D I P L D T L G L D V P D G	395	
1145	CTG CGG GAG TGG CGG CAC CGG GAC CGC GAA CTC CTC GCC GGC GTC AGC TCG TTC GGC ATG	1264	
396	L R E W P H P D R E L L A G V S S F G M	415	
1205	GGC GGC ACC AAC GCC CAC GTC GTC CTC AGC GAA GGC CCC GCG CAG GGC GGC GAG CAG CCC	1324	
416	G G T N A H V V L S E G P A Q G G E Q P	435	
1265	GCC ATC GAT GAG GAG ACC CCC GTC GAC AGC GGG GGC GCA CTG CCC TTC GTC GTC ACC GGC	1384	
436	G I D E E T P V D S G A A L P F V V T G	455	
1325	CGC GGC GGC GAG GCC CTG CGC GCC CAG CGG CGC CTG CAC GAG GGC GTC GAA CGG GAC	1444	
456	R G G E A L R A Q A R R L H E A V E A D	475	

Figure 23

1445	CCG GAG CTC GCG CCC GCC GCA CTC GCC CGG TCG CTG GTC ACC ACC CGT ACG GTC TTC ACG	1504
476	P E L A P A A L A R S L V T T R T V F T	495
1505	CAC CGG TCG GTC GTC CTC GCC CGG GAC CGC GCC CGG CTC CTC GAC GGC CTC GGC GCC CTC	1564
496	H R S V V L A P D R A R L L D G L G A L	515
1565	GCC GCC GGG ACG CCC CGC CCC GGC GTG GTC ACC GGC ACC CCC GCC CCC GGG CGC CTC GCC	1624
516	A A G T P A P G V V T G T P A P G R L A	535
1625	GTC CTG TTC AGC CGC CAG GGT GCC CAA CGT ACG GGC ATG GGC ATG GAG TTG TAC GCC GCC	1684
536	V L F S G Q G A Q R T G M G M E L Y A A	555
1685	CAC CCC GCC TTC CGC ACG GCC TTC GAC GCC GTC GCC GAA CTG GAC CCC CTC CTC GAC	1744
556	H P A F A T A F D A V A A E L D P L L D	575
1745	CGG CCC CTC GCC GAA CTC GTC CGC GCG GGC GAC ACC CTC GAC CGC ACC GTC CAC ACA CAG	1804
576	R P L A E L V A A G D T L D R T V H T Q	595
1805	CCC CGG CTC TTC GCC GTG GAG GTC GCC CTC CAC CGC CTC GTC GAG TCC TGG GGC GTC ACG	1864
596	P A L F A V E V A L H R L V E S W G V T	615
1865	CCC GAC CTG CTC GCC CGC CAC TCC GTC CGC GAG ATC ACG GCC GCC CAC GTC GCC GGG GTC	1924
616	P D L L A G H S V G E I S A A H V A G V	635
1925	CTG TCG CTG CGC GAC GCC CCC CTC GTC CGC CGC CGC CGC CTC ATG CAG CGC CTC	1984
636	L S L R D A A R L V A A R G R L M Q A L	655
1985	CCC GAG GGC GGC CGG ATG GTC CGC GAG GCG ACG GAG GAG GAA GTG CTT CCG CAC CTC	2044
656	P E G G A M V A V E A S E E E V L P H L	675
2045	GCG GGA CGC GAG CGG GAG CTC CTC CGC GTG AAC GGC CCC CGC CGG GTC GTC CTC	2104
676	A G R E R E L S L A A V N G P R A V V L	695
2105	GCG GGC GCC GAG CGC GCC GTC CTC GAC GTC GCC GAG CTG CTG CGC GAA CAG CGC CGC CGG	2164
696	A G A E R A V L D V A E L L L R E Q G R R	715
2165	ACG AAG CGG CTC AGC GTC TCG CAC GCC TTC CAC TCG CGG CTC ATG GAG CGG ATG CTC GAC	2224
716	T K R L S V S H A F H S P L M E P M L D	735
2225	GAC TTC CGC CGG GTC GTC GAA GAG CTG GAC TTC CAG GAG CCC CGC GTC GAC GTC GTG TCC	2284
736	D F R R V V E E L D F Q E P R V D V V S	755
2285	ACG GTG ACG GGC CTG CCT GTC ACA CGG GGC CAA TGG ACC GAT CCC GAG TAC TCC GTG GAC	2344
756	T V T G L P V T A G Q W T D P E Y W V D	775
2345	CAG GTC CGC AGG CCC GTA CGC TTC CTC GAC GCC GTA CGC ACC CTG GAG GAA TCG GGC GCC	2404
776	Q V R R P V R F L D A V R T L E E S G A	795
2405	GAC ACC TTC CTG GAG CTC GGT CCC GAC CGG GTC TGC TCC CGG ATG GCG GGC GAC TCC GTA	2464
796	D T F L E L G P D G V C S A M A A D S V	815
2465	CGC GAC CAG GAG GCC ACC CGG GTC TCC CGC CTG CGC AAG GGC CGC CGG GAG CCC CAG	2524
816	R D Q E A A T A V S A L R K G R P E P Q	835
2525	TCG CTG CTC GCC GCA CTC ACC ACC GTC TTC GTC CGG CGC CAC GAC GTC GAC TGG ACC GCC	2584
836	S L L A A L T T V F V R G H D V D W T A	855
2585	GCG CAC GGG AGC ACC GGC ACG GTC AGG GTG CCC CTG CGG ACC TAC GAC GCC TTC CAG CGC GAA	2644
856	A H G S T G T V R V P L P T Y A F Q R E	875
2645	CGC CAC TGG TTC GAC GGC GCC CGG CGA ACG GCG CGC CGG CTC ACG GCG CGC CGA TCG GGC	2704
876	R H W F D G A A R T A A P L T A G R S G	895
2705	ACC GGT CGG GGC ACC GGC CGG GCC CGG GGT GTG ACC TCG CGC GAG CGC GAG CGC GAG CGC	2764
896	T G A G T G P A A G V T S G E G E G E G	915
2765	GAG GGC GCG GGT CGG GGT GAT CGG CGG GCT CGC CAC GAG ACG ACC GAG CGC CGC GTG	2824
916	E G A G A G G G D R P A R H E T T E R V	935
2825	CGC GCA CAC GTC GCC GTC CTC GAG TAC GAC GAC CGG ACC CGC GTC GAA CTC CGC CTC	2884
936	R A H V A A V L E Y D D P T R V E L G L	955
2885	ACC TTC AAG GAG CTG CGC TTC GAC TCC ATG TCC GTC GAG CTG CGG AAC GCG CTC GTC	2944
956	T F K E L G F D S L M S V E L R N A L V	975
2945	GAC GAC ACG GGA CTG CGC CTG CCC AGC GGA CTG CTC TTC GAC CAC CGG ACG CGG CGC GCC	3004
976	D D T G L R L P S G L L F D H P T P R A	995

Figure 23 cont.

3005	CTC	GCC	GCC	CAC	CTG	GGC	GAC	CTG	CTC	ACC	GGC	GGC	GAG	ACC	GGA	TCG	GCC	GAC	3064	
996	L	A	A	H	L	G	D	L	L	T	G	G	S	G	E	T	G	S	A	1015
3065	GGG	ATA	CCG	CCC	GGG	ACC	CCG	GGG	GAC	ACC	ACC	GCC	GAG	CCC	ATC	GGG	ATC	ATC	GGC	3124
1016	G	I	P	P	A	T	P	A	D	T	T	A	E	P	I	A	I	I	G	1035
3125	GCC	TGC	GGC	TAC	CCC	GGC	GGC	GTC	ACC	TCC	CCC	GAG	GAC	CTG	TGG	CGG	CTC	GTC	GCC	3184
1036	A	C	R	Y	P	G	G	V	T	S	P	E	D	L	W	R	L	V	A	1055
3185	GGG	GGC	GAC	GCC	GTC	TCG	GGG	CTG	CCC	ACC	GAC	CGC	GGC	TGG	GAC	GAG	GAC	CTC	TTC	3244
1056	G	R	D	A	V	S	G	L	P	T	D	R	G	W	D	E	D	L	F	1075
3245	GCC	GAC	CCC	GAC	GGC	AGC	GGC	AAG	AGC	TCG	GTC	CGC	GAG	GGC	GGA	TTC	CTG	CAC	GAC	3304
1076	A	D	P	D	R	S	G	K	S	S	V	R	E	G	G	F	L	H	D	1095
3305	GCC	CTG	TTC	GAC	GCC	GGC	TTC	TTC	GGG	ATA	TCG	CCC	CGC	QAC	CCC	CTC	GCC	ATG	GAC	3364
1096	A	L	F	D	A	G	F	F	G	I	S	P	R	E	A	L	G	M	D	1115
3365	CAG	CAG	GGG	CTG	CTG	GAG	ACG	GCA	TCG	GAG	GGC	GTG	GAG	CGC	GCA	GGG	CTC	GAC	CCC	3424
1116	Q	Q	R	L	L	E	T	A	W	E	A	V	E	R	A	G	L	D	P	1135
3425	GAA	GGC	CTC	AAG	GGC	AGC	GGG	ACG	GGC	GTC	TTC	GGC	GGC	ACC	GCC	CTG	GAC	TAC	GGC	3484
1136	E	G	L	K	G	S	R	T	A	V	F	V	G	A	T	A	L	D	Y	1155
3485	CCG	GGC	ATG	CAC	GAC	GGC	GCC	GAG	GGC	GTC	GAG	GGC	CAC	CTC	CTG	ACC	GGG	ACC	ACG	3544
1156	P	R	M	H	D	G	A	E	G	V	E	G	H	L	L	T	G	T	T	1175
3545	AGC	GTG	ATG	TCG	GGC	GGC	ATC	GCC	TAC	CAG	CTC	GGC	CTC	ACC	GGT	CCT	GGG	GTC	ACC	3604
1176	S	V	M	S	G	R	I	A	Y	Q	L	G	L	T	G	P	A	V	T	1195
3605	GAC	ACG	GCC	TGC	TCG	TCC	TCG	CTC	GTC	GGG	CTG	CAC	CTG	GCC	GTC	CGT	TCG	CTG	GGG	3664
1196	D	T	A	C	S	S	S	L	V	A	L	H	L	A	V	R	S	L	R	1215
3665	GGC	GAG	TCG	AGC	CTC	GGC	CTC	GCC	GGG	ACC	GTC	ATG	TCG	ACA	CCG	GGC	ATG	TTC	3724	
1216	G	E	S	S	L	A	G	G	A	T	V	M	S	T	P	G	M	F	1235	
3725	GTC	GAG	TTC	TCG	GGG	CAG	GGC	GGC	CTC	GCC	GCC	GAC	GGC	TCC	AAG	GCC	TTC	TCC	GAC	3784
1236	V	E	F	S	R	Q	R	G	L	A	A	D	G	R	S	K	A	F	S	1255
3785	TCC	GCC	GAC	GGC	ACC	TCC	TGG	GCC	GAG	GGC	GTC	GGC	CTC	GTC	GTC	GAG	GGG	CTC	TCG	3844
1256	S	A	D	G	T	S	W	A	E	G	V	G	L	L	V	V	E	R	L	1275
3845	GAC	GCC	GAG	GGC	AAC	GGC	CAC	CCC	GTG	CTC	GGC	GTG	ATC	CGG	GGC	AGC	GGG	GTC	AAC	3904
1276	D	A	E	R	N	G	H	P	V	L	A	V	I	R	G	S	A	V	N	1295
3905	GAC	GGC	GCC	TCC	AAC	GGG	CTC	ACC	GCC	CCC	AAC	GGC	CCG	TOC	CAG	GGC	GTC	ATC	CGA	3964
1296	D	G	A	S	N	G	L	T	A	P	N	G	P	S	Q	Q	R	V	I	1315
3965	CAG	GCC	CTG	GCC	GAC	GGC	GGG	CTC	ACC	CGG	GCC	GAC	GTC	GGC	GTC	GAG	GGG	CAC	GGT	4024
1316	Q	A	L	A	D	A	G	L	T	P	A	D	V	D	A	V	E	A	H	1335
4025	ACG	GGT	ACC	CGG	CTC	GGC	GAC	CCC	ATC	GAG	GCC	GAG	GGC	ATC	CTC	GGC	ACC	TAC	GGC	4084
1336	T	G	T	R	L	G	D	P	I	E	A	E	A	I	L	G	T	Y	G	1355
4085	GAC	CGG	GGC	GAG	GGC	GCT	CTC	CAG	CTC	GGC	TCG	CTG	AAQ	TCG	AAC	ATC	GGC	CAC	GGG	4144
1356	D	R	G	E	G	A	P	L	Q	L	G	S	L	K	S	N	I	G	H	1375
4145	CAG	GCC	GCC	GCG	GGC	GTG	GGC	GGG	CTC	ATC	AAG	ATG	GTC	CTC	GGC	ATG	GGC	CAC	GGC	4204
1376	Q	A	A	A	G	V	G	G	L	I	K	M	V	L	A	M	R	H	G	1395
4205	CTG	CCC	AGG	ACG	CTC	CAC	GTG	GAC	CGG	CCC	ACC	ACC	CGC	GTC	GAC	TGG	GAG	GCC	GGC	4264
1396	L	P	R	T	L	H	V	D	R	P	T	T	R	V	D	W	E	A	G	1415
4265	GTC	GAG	CTC	CTC	ACC	GAG	GAG	GGG	GGG	GAG	GGC	GGC	GGC	GGC	GGC	GGC	GGG	GGC	GGG	4324
1416	V	E	L	L	T	E	E	R	E	W	P	E	T	G	R	P	R	R	A	1435
4325	ATC	TCC	TCC	TTC	GGC	ATC	AGC	GGC	ACC	AAC	GCC	CAC	ATC	GTG	GTC	GAA	CAG	GCC	CCG	4384
1436	I	S	S	F	G	I	S	G	T	N	A	H	I	V	V	E	Q	A	P	1455
4385	GCC	GGG	GAG	GGG	GGG	GTC	ACC	ACC	GCC	CCG	GAA	GCA	GGG	GAA	GCC	GGG	GAA	GCG	GCG	4444
1456	A	G	E	A	A	V	T	T	A	P	E	A	G	E	A	G	E	A	A	1475
4445	GAC	ACC	ACC	GCC	ACC	ACG	CGG	GGC	GTC	GGC	4504									
1476	D	T	T	A	T	T	T	P	A	A	V	G	V	P	E	P	V	R	A	1495
4505	GTC	GTG	GTC	TCC	GGC	GGG	GAC	GCC	GCC	CTG	GGC	GGC	CAG	GCC	GTC	GGG	CTG	GGG	ACC	4564
1496	V	V	V	S	A	R	D	A	A	A	L	R	A	Q	A	V	R	L	R	1515

Figure 23 cont.

4565 TTC CTC GAC GGC CGA CCG GAC GTC ACC GTC GCC GAC CTC GGA CGC TCG CTG GCC CCC CGT 4624
 1516 F L D G R P D V T V A D L G R S L A A R 1535
 4625 ACC GCC TTC GAG CAC AAG GCC GCC CTC ACC ACC GCC ACC AGG GAC GAG CTG CTC GCC GGG 4684
 1536 T A F E H K A A L T T A T R D E L L A G 1555
 4685 CTC GAC GCC CTC GGC CGC GGG GAG CAA CGC ACC GGC CTG GTC ACC GCC GAA CGG GCC AGG 4744
 1556 L D A L G R G E Q A T G L V T G E P A R 1575
 4745 GCC GGA CGC ACY GCC TTC CTG TTC ACC GCC CAG GGA CGC CAG CGC GTC GCC ATG GGC GAG 4804
 1576 A G R T A F L F T G Q G A Q R V A M G E 1595
 4805 GAA CTG CGC GCC CGC CAC CCC CTG TTC ACC GCC CTC GAC ACC GTG TAC GCG GCC CTC 4864
 1596 E L R A A H P V F A A A L D T V Y A A L 1615
 4865 GAC CGT CAC CTC GAC CGG CGG CTG CGG GAG ATC GTC ACC GCC GGC GGG GAG GAG CTG GAC CTC 4924
 1616 D R H L D R P L R E I V A A G E E L D L 1635
 4925 ACC CGG TAC ACC CAG CCC GCC CTC TTC GGC GAG GTG GCG CTG TTC CGC CTC CTC GAA 4984
 1636 T A Y T Q P A L F A F E V A L F R L L E 1655
 4985 CAC CAC GGC CTC GTC CCC GAC CTG CTC ACC GGC CAC TCC GTC GGC GAG ATC GCC GCC GCG 5044
 1656 H H G L V P D L L T G H S V G E I A A A 1675
 5045 CAC GTC GCC GGT GTC CTC TCC CTC GAC GAC GCA CGT CTC GTC ACC GCC CGC GGC CGG 5104
 1676 H V A G V L S L D D A A R L V T A R G R 1695
 5105 CTC ATG CAG TCG GCC CGC GAG GGC GCG ATG ATC GCC GTG CAG GCG GGC GAG GCC GAG 5164
 1696 L M Q S A R E G G A M I A V Q A G E A E 1715
 5165 GTC GTC GAG TCC CTG AAG GCC TAC GAG GGC AGG GTC GCC GTC GCC GCC GTC AAC GGA CCC 5224
 1716 V V E S L K G Y E G R V A V A A V N G P 1735
 5225 ACC GCC GTG GTC GTC TCC GGC GAC GCG GAC GCC GGC GAG GAG ATC CGC CCC GTA TGG GCG 5284
 1736 T A V V V S G D A D A A E E I R A V W A 1755
 5285 GGA CGC CGG CGG ACC CGC AGG CTG CGC GTC AGC CAC GAC TTC CAC TCC CCG CAC ATG 5344
 1756 G R G R R T R R L R V S H A F H S P H M 1775
 5345 GAC GAC GTC CTC GAC GAG TTC CTC CGG GTC GCC GAG GGC CTG ACC TTC GAG GAG CGG CGG 5404
 1776 D D V L D E F L R V A E G L T F E E P R 1795
 5405 ATC CCC GTC GTC TCC ACG GTC ACC GGC GCG CTC GTC AGC TCC GGC GAG CTC ACC TCG CCC 5464
 1796 I P V V S T V T G A L V T S G E L T S P 1815
 5465 CGG TAC TGG GTC GAC CAG ATC CGG CGG CCC GTG CGC TTC CTG GAC GCC GTC CGC ACC CTG 5524
 1816 A Y W V D Q I R R P V R F L D A V R T L 1835
 5525 GCC GCC CAG GAC CGG ACC GTC CTC GTC GAG ATC GGC CCC GAC GCC GTC CTC ACG GCA CTC 5584
 1836 A A Q D A T V L V E I G P D A V L T A L 1855
 5585 GCC GAG GAG GCT CTC GCG CCC GGC ACG GAC GCC CGG GAC GCG GAC GTC ACG GTC GTC 5644
 1856 A E E A L A P G T D A P D A R D V T V V 1875
 5645 CGG CTG CTG CGC CGG CCC GAG CCC GAG ACC CTC GTC GCC GGC GGT CTC GCG ACC GCC 5704
 1876 P L L R A G R P E P E T L A A G L A T A 1895
 5705 CAT GTC CAC GGC GCA CCC TTG GAC CGG CGG TCG TTC CTC CGG GAC GGG CGG CGC ACG GAC 5764
 1896 H V H G A P L D R A S F F P D G R R T D 1915
 5765 CTG CCC ACG TAC GCC TTC CGG CGC GAG CAC TAC TGG CTG AGC CCC GAG GCC CGT ACG GAC 5824
 1916 L P T Y A F R R E H Y W L T P E A R T D 1935
 5825 GCC CGC GCA CTC GGC TTC GAC CGG CGG CGC ACC CGG CTG CTG AGC ACC ACG GTC GAG GTC 5884
 1936 A R A L G F D P A R H P L L T T T V E V 1955
 5885 GCC GGC GGC GAC CGC GTC CTG ACC GGC CGT CTC TCC CTG ACC GAC CAG CCC TGG CTG 5944
 1956 A G G D G V L L T G R L S L T D Q P W L 1975
 5945 GCC GAC CAC ATG GTC AAC GGC GCC GTC CTG TTG CGG GCC ACC GCC TTC CTG GAG CTC GGC 6004
 1976 A D H M V N G A V L L P A T A F L E L A 1995
 6005 CTC GCG CGG CGC GAC CAC GTC GGG CGG GTC CGG GTG GAG GAA CTC ACC CTC GAA CGG CGG 6064
 1996 L A A G D H V G A V R V E E L T L E A P 2015
 6065 CTC GTC CTG CCC GAG CGG CGC GCC GTC CGC ATC CAG GTC CGC GTG AGC GGC GAC CGC GAG 6124
 2016 L V L P E R G A V R I Q V G V S G D G E 2035

Figure 23 Cont.

6125 TCG CGG GCC CGG CGC ACC TTC GGT GTG TAC AGC ACC CCC GAC TCC GGC GAC ACC GGT GAC	6184
2036 S P A G R T F G V Y S T P D S G D T G D	2055
6185 GAC GCG CCC CGG GAG TGG ACC CGC CAT GTC TCC GGC GTA CTC GGC GAA GGG GAC CCG GCG	6244
2056 D A P R E W T R H V S G V L G E G D P A	2075
6245 ACG GAG TCG GAC CAC CCC GGC ACC GAC GGG GAC GGT TCA GCG CCC TGG CGG CCT GCG GCG	6304
2076 T E S D H P G T D G D G S A A W P P A A	2095
6305 GCG ACC GCC ACA CCC CTC GAC GGC GTC TAC GAC CGG CTC GCG GAG CTC GGC TAC GGA TAC	6364
2096 A T A T P L D G V Y D R L A E L G Y G Y	2115
6365 GGT CGG GCC TTC CAG GGC CTG ACG GGG CTG TGG CGG GAC GGC GCC GAC ACG CTC GCC GAG	6424
2116 G P A F Q G L T G L W R D G A D T L A E	2135
6425 ATC CGG CTG CCC GCG CAG CAC GAG AGC GGG GGG CTC TTC GGC GTA CAC CGG GCG CTG	6484
2136 I R L P A A Q H E S A G L F G V H P A L	2155
6485 CTC GAC GCG GCG CTC CAC CGG ATC GTC CTG GAG GGC AAC TCA GCT GCC GGT GCC TGT GAC	6544
2156 L D A A L H P I V L E G N S A A G A C D	2175
6545 GCC GAT ACC GAC GCG ACC GAC CGG ATC CGG CTG CGG TTC GCG TGG CGG GGG GTG ACC CTC	6604
2176 A D T D A T D R I R L P F A W A G V T L	2195
6605 CAC GCC GAA GGG GCC ACC GCG CTC CGC GTA CGG ATC ACA CCC ACC GGC CGG GAC ACG GTC	6664
2196 H A E G A T A L R V R I T P T G P D T V	2215
6665 ACG CTC CGC CTC ACC GAC ACC GGT GCG CCC GTG GCC ACC GTG GAG TCC CTG ACC CTG	6724
2216 T L R L T D T T G A P V A T V E S L T L	2235
6725 CGC CGG GTG GCG AAG GAC CGG CTG GGC ACC ACC GCG GGG CGC GTC GAC GAC GCC CTG TTC	6784
2236 R A V A K D R L G T T A G R V D D A L F	2255
6785 ACG GTC GTG TGG ACG GAG ACC GGC ACA CGG GAA CCC GCA GGG CGC GGA GCC GTG GAG GTC	6844
2256 T V V W T E T G T P E P A G R G A V E V	2275
6845 GAG GAA CTC GTC GAC CTC GCC GGC CTC GGC GAC CTC GTG GAG CTC GGC GCC GCG GAC GTC	6904
2276 E E L V D L A G L G D L V E L G A A D V	2295
6905 GTC CTC CGG GCC GAC CGG TCG ACG CTC GAC GGG GAC CGG TCC GGC GCC CGC CGC ACA GCC	6964
2296 V L R A D R W T L D G D P S A A A R T A	2315
6965 GTC CGG CGC ACC CTC GCC ATC GTC CAG GAG TTC CTG TCC GAG CGG CGC TTC GAC GGC TCG	7024
2316 V R R T L A I V Q E F L S E P R F D G S	2335
7025 CGA CTG GTG TGC GTC ACC AGG GGC GCG GTC GCC GCA CTC CCC GGC GAG GAC GTC ACC TCC	7084
2336 R L V C V T R G A V A A L P G E D V T S	2355
7085 CTC GCC ACC GGC CCC CTC TGG GGC CTC GTC CGC TCC GGC CAG TCC GAG AAC CGG GGA CGC	7144
2356 L A T G P L W G L V R S A Q S E N P G R	2375
7145 CTG TTC CTC CTG GAC CTG GGT GAA CGC GAA GGC GAG CGC GAC CGG GCA GAG GAG CTG ATC	7204
2376 L F L L D L G E G E G E R D G A E E L I	2395
7205 CGC GCG GCC ACC GGC GGG GAC GAG CGG CAG CTC CGC GCA CGG GAC GGC CGA CTG CTC CGC	7264
2396 R A A T A G D E P Q L A A R D G R L L A	2415
7265 CGG AGG CTG GCC CGT ACC GCC GCC CTT TCG AGT GAG GAC ACC GCC GGC CGC GCC GAC CGT	7324
2416 P R L A R T A A L S S E D T A G G A D R	2435
7325 TTC GGC CCC GAC GGC ACC GTC CTC GTC ACC GGG GGC ACC GGA GGC CTC GGA GCG CTC CTC	7384
2436 F G P D G T V L V T G G T G G L G A L L	2455
7385 GCC CGC CAC CTC GTG GAG CGT CAC GGG GTG CGC CGG CTG CTG CTG GTG AGC CGC CGC CGG	7444
2456 A R H L V E R H G V R R L L V S R R G	2475
7445 GCC GAC GCC CCC GGC GCG GCC GAC CTG CGC GAG GAC CTC CGG GGC CTC CGC GCG GAG GTG	7504
2476 A D A P G A A D L G E D L A G L G A E V	2495
7505 GCG TTC GCC GCC GGC GAC CGC GGC GAG AGC CTG GCG CGG CGC ATC GCC ACC GTG	7564
2496 A F A A A D A A D R E S L A R A I A T V	2515
7565 CCC GCC GAG CAT CGG CTG ACG GGC GTC GTG CAC AGC GCG GGA GTC GTC GAC GAC GCG ACG	7624
2516 P A E H P L T A V V H T A G V V D D A T	2535
7625 GTG GAG CGG CTC ACA CGG GAA CGG CTG GAC GCG GTA CTG CGC CGG AAG GTC GAC GAC GCG	7684
2536 V E A L T P E R L D A V L R P K V D A A	2555

Figure 23 cont.

7685 TGG AAC CTG CAC GAG CTC ACC AAG GAC CTG CGG CTC GAC GCC TTC GTC CTC TTC TCC TCC	7744
2556 W N L H E L T K D L R L D A F V L F S S	2575
7745 GTC TCC GGC ATC GTC GGC ACC GCC GGCG CAG GCC AAC TAC GCG GCG GCC AAC ACG GGC CTC	7804
2576 V S G I V G T A G Q A N Y A A A A N T G L	2595
7805 GAC GCC CTC GCC GCC CAC CGC GCC GCC ACG GGCG CTG GCC GCC ACG TCG CTG GCC TGG GGC	7864
2596 D A L A A H R A A T G L A A A T S L A W G	2615
7865 CTC TGG GAC GGC ACG CAC GGC ATG GGC GGC ACG CTC CGC GCC GAC CTC GCC CGC TGG	7924
2616 L W D G T H G M G G T L G A A D L A R W	2635
7925 AGC CGG GCC GGA ATC ACC CGG CTC ACC CGG CTG CAG GGC CTC CGC CTC TTC GAC GCC GCG	7984
2636 S R A G I T P L T P L Q G L A L F D A A	2655
7985 GTC GCC AGG GAC GAC GCC CTC CTC GTA CCC GCC GGG CTC CGT CCC ACC GCC CAC CGG CCC	8044
2656 V A R D D A L L V P A G L R P T A H R G	2675
8045 ACG GAC GGA CAG CCT CCT GCG CTG TGG CGC GGC CTC GTC CGG CGC CGC CGC CGT GCC	8104
2676 T D G Q P P A L W R G L V R A R P R R A	2695
8105 CGG CGG ACC GCC GCC GAG GCG GCG GAC ACG ACC GGC CGC TGG CTG AGC GGG CTC GCC GCA	8164
2696 A R T A A E A A D T T G G G W L S G L A A	2715
8165 CAG TCC CCC GAG GAG CGG CGC AGC ACA GCC GTC ACG CTC GTG AGC GGT GTC GTC GCG GAC	8224
2716 Q S P E E R R S T A V T L V T G V V A D	2735
8225 GTC CTC GGG CAC GCC GAC TCC GCC CGG GTC CGG GCG GAG CGG TCC TTC AAG GAC CTC CGC	8284
2736 V L G H A D S A A V G A E R S F K D L G	2755
8285 TTC GAC TCC CTG GCC GGG GTG GAG CTC CGC AAC CGG CTG AAC GCC ACC GGC CTG CGG	8344
2756 F D S L A G V E L R N R L N A A T G L R	2775
8345 CTC CCC CGG ACC ACG GTC TTC GAC CAT CCC TCG CGC GCC GGG CTC CGG TCC CAT CTC CTC	8404
2776 L P A T T V F D H P S P A A L A S H L L	2795
8405 GCC CAG GTG CCC GGG TTG AAG GAG GGG ACG GCG GCG ACC GCG ACC GTC GTG GCC GAG CGG	8464
2796 A Q V P G L K E G T A A T A T V V A E R	2815
8465 CGC GCT TCC TTC GGT GAC CGT GCG ACC GAC GAC GAT CGG ATC GCG ATC GTC GTC GGC ATG GCA	8524
2816 G A S F G D R A T D D D P I A I V G M A	2835
8525 TGC CGC TAT CCG GGT GTG TCG TCG CGG GAG GAC CTG TGG CGG CTG GTG GCC GAG CGG	8584
2836 C R Y P G G V S S P E D L W R L V A E G	2855
8585 ACG GAC CGG ATC AGC GAG TTC CCC GTC AAC CGC CGC TGG GAC CTG GAG AGC CTC TAC GAC	8644
2856 T D A I S E F P V N R G W D L E S L Y D	2875
8645 CGG GAT CCC GAG TCG AAG CGC ACC ACG TAC TGC CGG GAG GGC CGG GGG TTC CTG GAA GGC GCC	8704
2876 P D P E S K G T T Y C R E G G F L E G A	2895
8705 GGT GAC TTC GAC GCC TTC TTC GGC ATC TCG CGG CGC GAG GGC CGC CTG GTG ATG GAC CGG	8764
2896 G D F D A A F F G I S P R E A L V M D P	2915
8765 CAG CAG CGG CTG CTG CGG GTG TCC TGG GAG GGC CTG GAA CGC CGG CGC ATC GAC CGG	8824
2916 Q Q R L L E V S W E A L E R A G I D P	2935
8825 TCC TCG CTG CGC AGC CGC GGT GGT GTC TAC GTG GGC GCC CGC CAC CGC TCG TAC GAC	8884
2936 S S L R G S R G G V Y V G A A H G S Y A	2955
8885 TCC GAT CCC CGG CTG GTG CCC GAG GGC TCG GAG GGC TAT CTG CTG ACC CGC AGC GCC GAC	8944
2956 S D P R L V P E G S E G Y L L T G S A D	2975
8945 CGG GTG ATG TCC GGC CGC ATC TCC TAC GCG CTC GGT CTC GAA CGA CGG TCC ATG ACG GTG	9004
2976 A V M S G R I S Y A L G L E G P S M T V	2995
9005 GAG AGC GCC TGC TCC TCC TCG CTG GTG GCG CTG CAT CTG GCG GTC GAA CGG CGG CTG CGG CAC	9064
2996 E T A C S S S L V A L H L A V R A L R H	3015
9065 CGC GAG TGC CGG CTC CGC CTG GCG GGG GTG GCG GTG ATG GCG GAT CGG CGG CGG CGG TIC	9124
3016 G E C G L A L A G G V A V M A D P A A F	3035
9125 GTG GAG TTC TCC CGG CAG AAG GGG CTG GCC GAC CGC CGC TGC AAG CGG TTC TCG CGC	9184
3036 V E F S R Q K G L A A D G R C K A F S A	3055
9185 CGC GCC GAC CGC ACC CGC TGG GCC GAG GGC GTC CGC GTG CTC GTC CTG GAG CGG CGG TCG TCG	9244
3056 A A D G T G W A E G V G V L V L E R L S	3075

Figure 23 Cont.

9245 GAC GCG CGC CGC CGG GGG CAC ACG GTC CTC GGC CTG GTC ACC GGC ACC GCG GTC AAC CAG 3076 D A R R A G H T V L G L V T G T A V N Q	9304 3095
9305 GAC GGT GCC TCC AAC GGG CTG ACC GCG CCC AAC GGC CCA GCC CAG CAA CGC GTC ATC GCC 3096 D G A S N G L T A P N G P A Q Q R V I A	9364 3115
9365 GAG GCG CTC GCC GAC GCC GGG CTG TCC CCG GAG GAC GTG GAC GCG GTC GAG GCG CAC GGC 3116 E A L A D A G L S P E D V D A V E A H G	9424 3135
9425 ACC GGC ACC CGG CTC GCC GAC CCC ATC GAG GCC GGG GCG CTG CTC GCC GCC TCC GGA CGG 3136 T G T R L G D P I E A G A L L A A S G R	9484 3155
9485 AAC CGT TCC GGC GAC CAC CGG CTG TGG CTC GGC TCG CTG AAG TCC AAC ATC GGG CAT GCC 3156 N R S G D H P L W L G S L K S N I G H A	9544 3175
9545 CAG GCC GCC GGC GGT GTC GGC GGC ATC AAG ATG CTC CAG GCG CTG CGG CAC GGC TTG 3176 Q A A A G V G G V I K M L Q A L R H G L	9604 3195
9605 CTG CCC CGC ACC CTC CAC GCC GAC GAG CGG ACC CCG CAT GCC GAC TGG AGC TCC GGC CGG 3196 L P R T L H A D E P T P H A D W S S G R	9664 3215
9665 GTA CGG CTG CTC ACC TCC GAG GTG CCG TGG CAG CGG ACC GGC CGG CCC CGG CGG ACC GGG 3216 V R L L T S E V P W Q R T G R P R R T G	9724 3235
9725 GTG TCC GCC TTC GGC GTC GGC ACC AAT GCC CAT GTC GTC CTC GAA GAG GCA CCC GCC 3236 V S A F G V G G T N A H V V L E E A P A	9784 3255
9785 CCG CCC CGG CCG GAA CGG GCC GGG GAG GCC CCC GGC CGC TCC CGC GCC GCA GAA GGG CGG 3256 P P A P E P A G E A P G G S R A A E G A	9844 3275
9845 GAA GGG CCC CTG GCC TGG GTG GTC TCC GGA CGC GAC GAG CGG GCC CTG CGG TCC CAG GCC 3276 E G P L A W V V S G R D E P A L R S Q A	9904 3295
9905 CGG CGG CTC CCC GAC CAC CTC TCC CGC ACC CCC GGG GCC CGC CCG CGT GAC ATC GCC TTC 3296 R R L R D H L S R T P G A R P R D I A F	9964 3315
9965 TCC CTC GCC GCC ACG CGC GCA GCC TTT GAC CAC CGC GCC GTG CTG ATC GGC TCG GAC GGG 3316 S L A A T R A A F D H R A V L I G S D G	10024 3335
10025 GCC GAA CTC GCC GCC CTG GAC GCG TTG GCC GAA GGA CGC GAC GGT CGG GCG GTG GTG 3336 A E L A A A L D A L A E G R D G P A V V	10084 3355
10085 CGC GGA GTC CGC GAC CGG GAC GGC AGG ATG GCC TTC CTC TTC ACC GGG CAG CGC ACC CAG 3356 R G V R D R D G R M A F L F T G Q G S Q	10144 3375
10145 CGC GCC GGG ATG GCC CAC GAC CTG CAT GCC GCC CAT ACC TTC TTC GCG TCC GCC CTC GAC 3376 R A G M A H D L H A A H T F F A S A L D	10204 3395
10205 GAG GTG ACG GAC CGT CTC GAC CGG CTG CTC GGC CGG CGC CTC GGC GCG CTG CTG GAC GCC 3396 E V T D R L D P L L G R P L G A L L D A	10264 3415
10265 CGA CCC GCC TCG CCC GAA GCG GCA CTC CTG GAC CGG ACC GAG TAC ACC CAG CGG GCG CTC 3416 R P G S P E A A L L D R T E Y T Q P A L	10324 3435
10325 TTC GCC GTC GAG GTG GCG CTC CAC CGG CTG CTG GAG CAC TGG GGG ATG CGC CCC GAC CTG 3436 F A V E V A L H R L L E H W G M R P D L	10384 3455
10385 CTG CTG GGG CAC TCG GTG GGC GAA CTG GCG GCC CAC GTC GCG GGT GTG CTC GAT CTC 3456 L L G H S V G E L A A A H V A G V L D L	10444 3475
10445 CAC GAC GCC TGC GCG CTG GTG GCC CCC CGC GGC AGG CTG ATG CAG CGC CTG CGG CCC GGC 3476 D D A C A L V A A R G R L M Q R L P P G	10504 3495
10505 GGC GCG ATG GTC TCC GTG CGG GCC GGC GAG GAC GAG GTC CGC GCA CTG CTG GCC GGC CGC 3496 G A M V S V R A G E D E V R A L L A G R	10564 3515
10565 GAG GAC GCC GTC TGC GTC GCC GCG GTG AAC GGC CCC CGG TCG GTG GTG ATC TCC GGC GCG 3516 E D A V C V A A V N G P R S V V I S G A	10624 3535
10625 GAG GAA GCG GTG GCG GAG GCG GCG CAG CTC GCC GGA CGA CGC CGC CGC ACC AGG CGG 3536 E E A V A E A A Q L A G R G R R T R R	10684 3555
10685 CTC CGC GTC GCG CAC GCC TTC CAC TCA CCC CTG ATG GAC GCG ATG CTC GCC GGA TTC CGG 3556 L R V A H A F H S P L M D G M L A G F R	10744 3575
10745 GAG GTC GCC GCC CGC CTG CGC TAC CGG GAA CGG GAG CTG ACG GTC GTC TCC ACG GTC ACG 3576 E V A A G L R Y R E P E L T V V S T V T	10804 3595

Figure 23 cont.

10805 CGC CGG CCC GCC CGC CCC GGT GAA CTC ACC GGC CCC GAC TAC TGG GTG GCC CAG GTC CGT 10864 3596 G R P A R P G E L T G P D Y W V A Q V R 3615
10865 GAG CCC GTG CGC TTC GCG GAC GCG GTC CGC ACC GCA CAC CGC CTC GGA GCC CGC ACC TTC 10924 3616 E P V R F A D A V R T A H R L G A R T F 3635
10925 CTG GAG ACC GGC CGG GAC GGC GTG CTG TGC GGC ATG GCA GAG GAG TGC CTG GAG GAC GAC 10984 3636 L E T G P D G V L C G M A E E C L E D D 3655
10985 ACC GTG GCC CTG CTG CGG GCG ATC CAC AAG CCC GGC ACC GCG CCG CAC GGT CCG GCG GCT 11044 3656 T V A L L P A I H K P G T A P H G P A A 3675
11045 CCC CGC GCG CTG CGG GCG GCC GCC CGC CGG TAC GGC CGG GGC GCC CGG GTG GAC TGG GCC 11104 3676 P G A L R A A A A A Y G R G A R V D W A 3695
11105 GGG ATG CAC CCC GAC GGC CCC GAG GGG CGG GCC CGC CGC GTC GAA CTG CCC GTC CAC GCC 11164 3696 G M H A D G P E G P A R R V E L P V H A 3715
11165 TTC CCG CAC CGC CGC TAC TGG CTC GCC CGG CGC CGC GCG GAC ACC GAC GAC TGG ATG 11224 3716 F R H R R Y W L A P G R A A D T D D W M 3735
11225 TAC CGG ATC GGC TGG GAC CGG CTG CGG GCT GTG ACC GGC GGG GCC CGG ACC GCC CGC CGC 11284 3736 Y R I G W D R L P A V T G G A R T A G R 3755
11285 TGG CTG GTG ATC CAC CCC GAC AGC CGG CGC TGC CGG GAG CTG TCC GGC CAC GCC GAA CGC 11344 3756 W L V I H P D S P R C R E L S G H A E R 3775
11345 CGG CTG CGC GCC CGG CGC AGC CCC GTC CGG CTG CCC GTG GAC GCT CGG GCC CCC GAC 11404 3776 A L R A A G A S P V P L P V D A P A A D 3795
11405 CGG CGC TCC TTC GCG GCA CTG CTG CGC TCC GCC ACC GGA CCT GAC ACA CGA GGT GAC ACA 11464 3796 R A S F A A L L R S A T G P D T R G D T 3815
11465 CGC CGG CCC GTG GCC GGT GTG CTG TCG CTG TCC GAG GAG GAT CGG CCC CAT CGC CAG 11524 3816 A A P V A G V L S L L S E E D R P H R Q 3835
11525 CAC GCC CGG GTA CCC GCC GGG GTC CTG CGG ACG CTG TCC CTG ATG CAG GCT ATG GAG GAG 11584 3836 H A P V P A G V L A T L S L M Q A M E E 3855
11585 GAG CGG GTG GAG GCT CGC GTG TGG TGC GTC TCC CGC GCC CGG GTC GCC GCC GAC CGG 11644 3856 E A V E A R V W C V S R A A V A A A D R 3875
11645 GAA CGG CCC GTC CGC CGG CGC GCC CTG TGG CGG CTG CGG CGG GTG GCC GCC CTG GAA 11704 3876 E R P V G A A G A L W G L G R V A A L E 3895
11705 CGC CCC ACC CGG TGG CGC GGT CTC GTG GAC CTG CCC GCC TCG CCC GGT CGG GCG CAC TGG 11764 3896 R P T R W G G L V D L P A S P G A A H W 3915
11765 CGG CGC CGC GTG GAA CGG CTC GCC GGT CCC GAG GAC CAG ATC GGC GTG CGC CGG TCC CGC 11824 3916 A A A V E R L A G P E D Q I A V R A S G 3935
11825 AGT TGG GGC CGG CGC CTC ACC AGG CTG CGG CGC GAC GGC GGC CGC CGG ACG GGC GCA CCC 11884 3936 S W G R R L T R L P R D G G G R T A A P 3955
11885 CGC TAC CGG CGG CGC CGC ACG GTG CTC GTC ACC GGT CGC ACC GGC GCG CTC CGC GGG CAT 11944 3956 A Y R P R G T V L V T G G T G A L G G H 3975
11945 CTC GCC CGC TGG CTC GCC CGG CGC GGC GCA CAC CTG GCG CTC ACC AGC CGC CGG GGC 12004 3976 L A R W L A A A G A E H L A L T S R R G 3995
12005 CCG GAC CGG CCC CGC GCC GG A CTC GAG GGC GAA CTC CTC CTC CTG GGC GCC AAG GTG 12064 3996 P D A P G A A G L E A E L L L G A K V 4015
12065 ACG TTC GCC CGC TGC GAC ACC GCC GAC CGC GAC GGC CTC CGC CGG GTC CTG CGG GCG ATA 12124 4016 T F A A C D T A D R D G L A R V L R A I 4035
12125 CGG GAG GAC ACC CGG CTC ACC GCG GTG TTC CAC GCC CGG GGC GTC ACG CGC 12184 4036 P E D T P L T A V F H A A G V P Q V T P 4055
12185 CTG TCC CGT ACC TGG CCC GAG CAC TTC GCG GTC TAC GCG GGC AAG CGG CGG GGC GCC 12244 4056 L S R T S P E H F A D V Y A G K A A G A 4075
12245 GCG CAC CTG GAC GAA CTG ACC CGC GAA CTC GGC GGC GGA CTC GAC GCG TTC GTC CTC TAC 12304 4076 A H L D E L T R E L G A G L D A F V L Y 4095
12305 TCC TCC CGC CGC GGC GTC TGG CGC AGC GCC CGC CAG GGT GCC TAC GGC GCC AAC CGC 12364 4096 S S G A G V W G S A G Q G A Y A A N A 4115

Figure 23 cont.

12365 GCC CTG GAC GCG CTC GCC CGG CGC CGT GCG GCG GAC GGA CTC CCC GCC ACC TCC ATC GCC 12424
 4116 A L D A L A R R R A A D G L P A T S I A 4135
 12425 TGG GCC GTG TGG GGC GGC GGC GGT ATG GGG GCC GAC GAG GCG GGC GCG GAG TAT CTG GGC 12484
 4136 W G V W G G G G M G A D E A G A E Y L G 4155
 12485 CGG CGC GGT ATG CGC CCC ATG GCA CGG GTC TCC GCG CTC CGG GCG ATG GCC ACC GCC ATC 12544
 4156 R R G M R P M A P V S A L R A M A T A I 4175
 12545 GCC TCC GGG GAA CCC TGC CCC ACC GTC ACC CAC ACC GAC TGG GAG CCC TTC GGC GAG GGC 12604
 4176 A S G E P C P T V T H T D W E R F G E G 4195
 12605 TTC ACC GCC TTC CGG CCC AGC CCT CTG ATC GCG GGG CTC GGC ACG CGG GGC GGC CGG 12664
 4196 F T A F R P S P L I A G L G T P G G G R 4215
 12665 CGG CGG GAG ACC CCC GAG GAG GGG AAC GCC ACC GCT GCG GCG GAC CTC ACC GCC CTG CGG 12724
 4216 A A E T P E E G N A T A A A A D L T A L P 4235
 12725 CCC GCC GAA CTC CGC ACC GCG CTG CGC GAG CTG GTG CGA GCC CGG ACC GCC GCG CGG CTC 12784
 4236 P A E L R T A L R E L V R A R T A A A L 4255
 12785 GCC CTC GAC GAC CGG GCC GAG GTC GCC GAG GGC GAA CGG TTC CCC GCC ACC ATG GGC TTC GAC 12844
 4256 G L D D P A E V A E G E R F P A M G F D 4275
 12845 TCC CTG GCC ACC GTA CGG CTG CGC CGC GGA CTC GCC TCG GCC ACG GGC CTC GAC CTG CCC 12904
 4276 S L A T V R L R R G L A S A T G L D L P 4295
 12905 CCC GAT CTG CTC TTC GAC CGG GAC ACC CGG GCC GCG CTC GCC CCC CAC CTG GCC GAA CTG 12964
 4296 P D L L F D R D T P A A L A A H L A E L 4315
 12965 CTC GCC ACC GCA CGG GAC CAC GGA CCC CGC GGC CCC GGG ACC GGT GCC CGG CGG GCC GAT 13024
 4316 L A T A R D H G P G G P G T G A A P A D 4335
 13025 GCC GGA AGC GGC CTG CGG GCC CTC TAC CGG GAG GCC GTC CGC ACC GGC CGG GCC GCG GAA 13084
 4336 A G S G L P A L Y R E A V R T G R A A E 4355
 13085 ATG GCC GAA CTG CTC GCC GCT TCC CGG TTC CGC CCC GCC TTC GGG ACG CGG GAC CGG 13144
 4356 M A E L L A A A S R F R P A F G T A D R 4375
 13145 CAG CGG GTG GCC CTC GTG CGG CTG GCC GAC GGC GCG GAG GAC ACC GGG CTC CGG CTG CTC 13204
 4376 Q P V A L V P L A D G A E D T G L P L L 4395
 13205 GTG GGC TGC GCC GGG ACG GCG GTG GCC TCC GGC CGG GTG GAG TTC ACC GCC TTC GCC GGA 13264
 4396 V G C A G T A V A S G P V E F T A F A G 4415
 13265 GCG CTG CGG GAC CTC CGG CGG GCG GCC CGG ATG GCC GCG CTG CGG CAG CCC GGC TTT CTG 13324
 4416 A L A D L P A A A P M A A L P Q P G F L 4435
 13325 CGG GGA GAA CGA GTC CGG GCC ACC CGG GAG GCA TIG TTC GAG GCC CAG CGG GAA CGG CTG 13384
 4436 P G E R V P A T P E A L F E A Q A E A L 4455
 13385 CTG CGC TAC CGG GCC CGG CCC TTC GTG CTG CTG GGG CAC TCC GCC CGC GCC AAC ATG 13444
 4456 L R Y A A G R P F V L L G H S A G A N M 4475
 13445 GCC CAC GCC CTG ACC CGT CAT CTG GAG GCG AAC CGT GCG CCC GCA CGG CTG GTG CTC 13504
 4476 A H A L T R H L E A N G G G P A G L V L 4495
 13505 ATG GAC ATC TAC ACC CCC GCC GAC CCC CGC GCG ATG GGC GTC TGG CGG AAC GAC ATG TTC 13564
 4496 M D I Y T P A D P G A M G V W R N D M F 4515
 13565 CAG TGG GTC TGG CGG CGC TCG GAC ATC CCC CGG GAC GAC CAC CGC CTC ACG GCC ATG GGC 13624
 4516 Q W V W R R S D I P P D D H R L T A M G 4535
 13625 GCC TAC CAC CGG CTG CTT CTC GAC TGG TCG CCC ACC CCC GTC CGC GCC CCC GTC CTG CAT 13684
 4536 A Y H R L L D W S P T P V R A P V L H 4555
 13685 CTG CGC GCC CGG GAA CCC ATG GGC GAC TGG CCA CCC CGG GAC ACC CGC TGG CAG TCC CAC 13744
 4556 L R A A E P M G D W P P G D T G W Q S H 4575
 13745 TGG GAC GGC GCG CAC ACC ACC GCC ATC CCC CGG AAC CAC TTC ACG ATG ATG ACC GAA 13804
 4576 W D G A H T T A G I P G N H F T M M T E 4595
 13805 CAC GCC TCC GCC CGC CGG CTC GTG CAC GGC TGG CTC GCG GAA CGG ACC CGG TCC GGG 13864
 4596 H A S A A A R L V H G W L A E R T P S G 4615
 13865 CAG CGC GGG TCA CGG TCC CGC CGG GGG AGA GAG GAG AGG CGG TGA ACACGGCAGCCGGCCC 13928
 4616 Q G G S P S R A A G R E E R P * 4631

Figure 23 cont.

32/33

13929	GACCGGCACCGCCGGCGGCCACCAACGGCCCGGGCACAGGACCTGTCCCCGGGGACCGCAGGCTCCAATCA	14008
14009	CCCCGGGGCACAGTGGTTGCGCCGGCAACCAGGGAGACCCCTACGGG	14079
1	M I L R A G T A 8	
14080	GAC CGG GCA CCG TAC GAG GAA GAG ATC CCC GGG TAC CGA GCT CGA ATT CTT AAT TAA GGAG	14140
9 D P A P Y E E E I P G Y R A R I L N *		27
14141	GTCGTAG ATG AGT AAC AAG AAC GAT GAG CTG CAG CGG CAG GCC TCG GAA AAC ACC CTG	14201
1 M S N K N N D E L Q R Q A S E N T L		18
14202	GGG CTG AAC CCG GTC ATC GGT ATC CGC CGC AAA GAC CTG TTG AGC TCG GCA CGC ACC GTG	14261
19 G L N P V I G I R R K D L L S S A R T V		38
14262	CAG CGC CAG CCC CAA CGG CTG CAC AGC GCC AAG CAT GTG GCC CAC TTT GGC CTG	14321
39 L R Q A V R Q P L H S A K H V A H F G L		56
14322	GAG CTG AAG AAC GTG CTG CTG GGC AAG TCC AGC CTT GCC CCG GAA AGC GAC GAC CGT CGC	14381
59 E L K N V L L G K S S L A P E S D D R R		78
14382	TTC AAT GAC CCG GCA TGG AGC AAC CCA CTT TAC CGC CGC TAC CTG CAA ACC TAT CTG	14441
79 F N D P A W S N N P L Y R R Y L Q T Y L		98
14442	GCC TGG CGC AAG GAG CTG CAG GAC TGG ATC GGC AAC AGC GAC CTG TCG CCC CAG GAC ATC	14501
99 A W R K E L Q D W I G N S D L S P Q D I		118
14502	ACC CGC CGC CAG TTC GTC ATC AAC CTG ATG ACC GAA GCC ATG GCT CCG ACC AAC ACC CTG	14561
119 S R G Q F V I N L M T E A M A P T N T L		138
14562	TCC AAC CGG GCA GCA GTC AAA CGC TTC TTC GAA ACC GGC GGC AAG AGC CTG CTC GAT GGC	14621
139 S N P A A V K R F F E T G G K S L L D G		158
14622	CTG TCC AAC CTG GCC AAG GAC CTG GTC AAC AAC GGT GGC ATG CCC AGC CAG GTG AAC ATG	14681
159 L S N L A K D L V N N G G M P S Q V N M		178
14682	GAC GCC TTC GAG GTG GGC AAG AAC CTG GGC ACC ACT GAA GGC GCC CTG GTG TAC CGC AAC	14741
179 D A F E V G K N L G T S E G A V V Y R N		198
14742	GAT CTG CTG GAG CTG ATC CAG TAC AAG CCC ATC ACC GAG CAG GTG CAT GCC CGC CCG CTG	14801
199 D V L E L I Q Y K P I T E Q V H A R P L		218
14802	CTG GTG GTG CCG CAG ATC AAC AAG TTC TAC GTC TTC GAC CTG AGC CCG GAA AAG AGC	14861
219 L V V P P Q I N K F Y V F D L S P E K S		238
14862	CTG GCA CGC TAC TGC CTG CGC TCG CAG CAG CAG ACC TTC ATC ATC AGC TGG CGC AAC CCG	14921
239 L A R Y C L R S Q Q Q T F I I S W R N P		258
14922	ACC AAA GCC CAG CGC GAA TGG GGC CTG TCC ACC TAC ATC ATC GAC GCG CTC AAG GAG GCG GTG	14981
259 T K A Q R E W G L S T Y I D A L K E A V		278
14982	GAC GCG GTG CTG GCG ATT ACC GCC AGC AAG GAC CTG AAC ATG CTC GGT GCC TGC TCC GGC	15041
279 D A V L A I T G S K D L N M L G A C S G		298
15042	GCG ATC ACC TGC ACG GCA TTG GTC GGC CAC TAT GCC GGC CTC GGC GAA AAC AAG GTC AAT	15101
299 G I T C T A L V G H Y A A L G E N K V N		318
15102	GCC CTG ACC CTG CTG GTC AGC GTG CTG GAC ACC ACC ATG GAC AAC CAG GTC GCC CTG TTC	15161
319 A L T L L V S V L D T T M D N Q V A L F		338
15162	GTC GAC GAG CAG ACT TTG GAG GGC GCC AAG CGC CAC TCC TAC CAG GCC GGT GTG CTC GAA	15221
339 V D E Q T L E A A K R H S Y Q A G V L E		358
15222	GCG AGC GAG ATG GCC AAG GTG TTC GCC TGG ATG CGC CCC AAC GAC CTG ATC TGG AAC TAC	15281
359 G S E M A K V F A W M R P N D L I W N Y		378
15282	TGG GTC AAC AAC TAC CTG CTC GGC AAC GAG CGG CGG GTG TTC GAC ATC CTG TTC TGG AAC	15341
379 W V N N Y L L G N E P P V F D I L F W N		398
15342	AAC GAC ACC ACG CGC CTG CGG GCC TTC CAC GGC GAC CTG ATC GAA ATG TTC AAG AGC	15401
399 N D T T R L P A A F H G D L I E M F K S		418
15402	AAC CGG CTG ACC CGC CGG GAC GCC CTG GAG GTT TGC GGC ACT CCG ATC GAC CTG AAA CAG	15461
419 N P L T R P D A L E V C G T P I D L K Q		438
15462	GTC AAA TGC GAC ATC TAC AGC CTT GCC GGC ACC AAC GAC CAC ATC ACC CCG TGG CAG TCA	15521
439 V K C D I Y S L A G T N D H I T P W Q S		458

Figure 23 cont.

15522 TGC TAC CGC TCG GCG CAC CTG TTC GGC AAG ATC GAG TTC GTG CTG TCC AAC AGC GGC 459 C Y R S A H L F G G K I E F V L S N S G	15581 478
15582 CAC ATC CAG AGC ATC CTC AAC CCG CCA GGC AAC CCC AAG GCG CGC TTC ATG ACC GGT GCC 479 H I Q S I L N P P G N P K A R F M T G A	15641 498
15642 GAT CGC CCG GGT GAC CCG GTG GCC TGG CAG GAA AAC GCC ACC AAG CAT GCC GAC TCC TGG 499 D R P G D P V A W Q E N A T K H A D S W	15701 518
15702 TGG CTG CAC TGG CAA AGC TGG CTG GCC GAG CGT GCC GGC GAG CTG GAA AAG GCG CCG ACC 519 W L H W Q S W L G E R A G E L E K A P T	15761 538
15762 CGC CTG GGC AAC CGT GCC TAT GCC GCT GGC GAG GCA TCC CCG GGC ACC TAC GTT CAC GAG. 539 R L G N R A Y A A G E A S P G T Y V H E	15821 558
15822 CGT TGA GCTGCAGGGCCGTGGCACCTGGGGACGCCACGGTGTTGAATT 559 R *	15872 560

Figure 23 cont.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/20119

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C12P 7/62; C12N 9/88, 15/63; C07H 21/04
US CL :435/135, 232, 320.1; 536/23.2

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/135, 232, 320.1; 536/23.2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	US 5,371,002 A (DENNIS ET AL.) 06 December 1994, column 9, line 65 through column 10, line 62 and Figure 9.	4, 5, 22, 25-27, 33-35, 52, 57, 59 ----- 1-3, 6-21, 28, 29, 32, 38, 51
Y	FRASER, M.J. The Baculovirus-Infected Insect Cell as a Eukaryotic Gene Expression System. Curr. Top. Microbiol. Immunol. August 1992, Vol. 158, pages 131-172, especially pages 145-158.	1-3, 6-21, 38

Further documents are listed in the continuation of Box C.

See patent family annex.

• Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
E earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
O document referring to an oral disclosure, use, exhibition or other means		
P document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

03 APRIL 1997

Date of mailing of the international search report

15 APR 1997

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703) 305-3230

Authorized officer

KAWAI LAU

Telephone No.

(703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/20119

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JOSHI, A.N. et al. Construction, Expression, and Characterization of a Mutated Animal Fatty Acid Synthase Deficient in the Dehydrase Function. J. Biol. Chem. 25 October 1993, Vol. 268, No. 30, pages 22508-22513, see whole document.	13-21, 28, 29-32, 40-49, 51
X	US 5,229,279 A (PEOPLES ET AL) 20 July 1993, column 15, lines 43-68.	37, 58
Y	DONADIO, S. et al. Organization of the Enzymatic Domains in the Multifunctional Polyketide Synthase Involved in Erythromycin Formation in <i>Saccharopolyspora erythraea</i> . Gene. 01 February 1992, Vol. 111, No. 1, pages 51-60, see whole document.	30, 31, 40-49
Y	BEVITT, D.J. et al. 6-Deoxyerythronolide-B Synthase 2 from <i>Saccharopolyspora erythraea</i> . Eur. J. Biochem. February 1992, Vol. 204, No. 1, pages 39-49, see whole document.	30, 31, 40-49
X	Database Medline on Dialog, US National Library of Medicine, (Bethesda, MD, USA), No. 95187308, HAN, L. et al. 'Cloning and Characterization of Polyketide Synthase Genes for Jadomycin B Biosynthesis in <i>Streptomyces venezuelae</i> ISP5230', abstract, Microbiol., Vol. 140, No. 12, pages 3379-89, December 1994, see entire abstract.	53

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/20119

B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

APS; STN-indices Bioscience Patents
search terms: PHA, PHB, polyhydroxyalkanoate, polyhydroxybutyrate, polyketide, fatty acid synthase, dehydrase,
Alcaligenes eutrophus, Saccharopolyspora erythraea, S. venezuelae

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER: _____**

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.